A Workflow for Extracting Sanskrit Text from Transcripts and Translating it to English

[1]Meher Pranav Kurapati, [2]Harshith Reddy Soda,[3]Hemanth Chowdary Vallabhaneni,[4]Murali Mohan Vutukuru,[5]Lohith Reddy Burramukku,[6]Satish Thatavarthi
[1]Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
[2]Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
[3]Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
[4]Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
[5]Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
[6]Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India

Abstract

Introduction: The digitalization of historical and precious writings has gained prominence in recent years, primarily for the protection and accessibility of cultural property. However, extracting Sanskrit text from scanned photos poses significant challenges due to the intricate and diverse nature of Sanskrit script.

Objectives: This research aims to develop a novel methodology for extracting Sanskrit text from scanned photos using a hybrid approach that integrates Convolutional Neural Networks (CNN) and Optical Character Recognition (OCR) techniques. The study seeks to address the complexities of Sanskrit script by combining deep learning with classical text recognition methods.

Methods: The process involves pre-processing scanned images to enhance text visibility, followed by training a CNN model to detect and localize text regions accurately. Subsequently, an OCR algorithm, fine-tuned specifically for Sanskrit characters, is applied to the localized areas to ensure accurate reproduction of the text. This hybrid CNN-OCR model surpasses traditional OCR methods by effectively handling variations in font styles, sizes, and script complexities unique to Sanskrit.

Results: Experimental results demonstrate the efficacy of the proposed approach across a broad spectrum of Sanskrit materials, including ancient manuscripts and printed texts. The CNN-OCR hybrid model automates the text extraction process and significantly improves accuracy in identifying Sanskrit characters compared to conventional methods. This advancement contributes to the digital preservation and enhanced accessibility of Sanskrit literature and heritage.

Conclusions: The combination of CNN-based text identification and Sanskrit-tailored OCR offers a promising avenue for advancements in historical text digitization. By bridging modern technology with classical language studies, this study facilitates the preservation and dissemination of Sanskrit heritage, thus enriching cultural knowledge and understanding.

Keywords: Digitalization, Sanskrit Text Extraction, Convolutional Neural Networks, Optical Character Recognition, Deep Learning, Text Localization

1. Introduction

Sanskrit, often revered as the "language of the gods," stands as an ancient and profoundly significant classical language that has bequeathed humanity with an extensive repository of wisdom spanning literature, philosophy, science, and spirituality. The corpus of Sanskrit texts, stretching back millennia, remains a source of inspiration and scholarly pursuit. Yet, the challenge lies in accessing and studying these texts due to their often-archaic formats, many of which exist as scanned images within PDF documents. Extracting and translating this invaluable content poses a formidable obstacle for researchers and scholars worldwide.

This research article endeavours to introduce a novel and comprehensive workflow explicitly crafted to surmount this challenge, thereby facilitating access to Sanskrit texts for academic and research purposes. Our innovative workflow harnesses cutting-edge technology, incorporating Optical Character Recognition (OCR) and machine translation, to bridge the linguistic gap and unlock the treasures concealed within these scanned PDF images.

The primary objective of our workflow is to empower scholars and researchers to extract Sanskrit text from PDF images and translate it into English, thereby making this ancient knowledge accessible to a broader audience. In doing so, we aspire to catalyze interdisciplinary research, foster cross-cultural understanding, and preserve the cultural heritage embedded in Sanskrit literature.

The subsequent sections delineate the various steps of our workflow, encompassing preprocessing PDF images, extracting Sanskrit text using OCR tools meticulously trained for Sanskrit scripts, translating the text into English while maintaining its contextual integrity, and conducting post-translation quality assessments to ensure accuracy. Through these meticulous steps, we present a systematic and reliable approach for unlocking the wisdom encapsulated within Sanskrit texts, transcending language barriers, and paving the way for meaningful research and scholarship.

This workflow not only serves as a practical tool for researchers but also contributes significantly to broader digitization and preservation efforts for Sanskrit literature. It signifies a bridge between ancient wisdom and modern scholarship, facilitating the dissemination of knowledge and the perpetuation of the rich cultural heritage encapsulated within the Sanskrit language. As such, this innovative approach represents a pivotal stride towards bridging the temporal and linguistic gap, fostering a seamless continuity in the exploration of Sanskrit's timeless wisdom.

2. Literature Review

Sanskrit, often hailed as the "language of the gods," stands as a classical language rich in historical and cultural significance. The vast repository of knowledge encapsulated in Sanskrit texts spans diverse domains, making it a subject of enduring scholarly interest. However, the digitization of these texts, especially the extraction and translation processes, presents multifaceted challenges. This literature survey explores existing research and methodologies relevant to Sanskrit text extraction and translation, aiming to position the proposed workflow within the broader academic context.

The intricate nature of Sanskrit script, characterized by its complex morphology and varied writing styles, poses challenges for accurate text extraction. Scholars have grappled with the need for advanced technologies to navigate through these intricacies. Optical Character Recognition (OCR) and Convolutional Neural Networks (CNN) have emerged as key technologies in addressing the nuances of Sanskrit script, enhancing the precision of text extraction processes.

The preservation and accessibility of historical texts have been central concerns in academia. Previous research has emphasized the significance of leveraging technology for the digitization of ancient manuscripts. Sanskrit, with its rich literary heritage, is no exception. The literature survey reveals a growing interest in workflows that cater specifically to the challenges posed by Sanskrit scripts, contributing to the broader trend of digitizing historical texts.

An integral aspect of Sanskrit text extraction is the subsequent translation into English. Machine translation techniques have been explored to bridge the linguistic gap between classical languages and modern understanding. Existing studies recognize the importance of providing access to Sanskrit knowledge for a wider audience and emphasize the role of machine translation in achieving this goal.

The convergence of interdisciplinary studies, combining linguistic research with advancements in computer science, is evident in the literature. The proposed workflow aligns with these trends, recognizing the symbiotic relationship between technological innovations and the preservation of cultural heritage embedded in Sanskrit literature. This interdisciplinary approach positions the workflow as a holistic solution for both scholars and enthusiasts alike.

In conclusion, the literature survey provides a comprehensive overview of the challenges and opportunities in Sanskrit text extraction and translation. By integrating advanced technologies, addressing script intricacies, and contributing to the broader digitization efforts, the proposed workflow aligns with the evolving landscape of scholarly endeavors in preserving and making accessible the timeless wisdom encapsulated in Sanskrit texts.

3. Methodology

The methodology for extracting Sanskrit text from transcripts and translating it to English involves a systematic and multi-step process. This workflow seamlessly integrates advanced technologies such as Optical Character Recognition (OCR), Convolutional Neural Networks (CNN), and machine translation to overcome the challenges posed by the intricate nature of Sanskrit script. The following paragraphs detail the key steps involved in this comprehensive methodology.

1. Preprocessing and Image Extraction:

- PDF Loading with FitZ:

The process initiates with the utilization of the FitZ library to load the PDF file containing the Sanskrit transcripts or scanned documents. FitZ provides functionalities to interact with PDF documents, facilitating the extraction of textual content and images.

- Page Processing:

Each page of the loaded PDF undergoes processing to extract relevant images. FitZ enables the extraction of images embedded within the PDF. This step ensures that the subsequent workflow operates on visual content relevant to Sanskrit text, preparing the groundwork for subsequent text extraction and analysis.

- Image Relevance Check:

During the extraction process, a check for image relevance is conducted. This involves assessing the content of each extracted image to confirm its alignment with the textual material. Non-relevant images, if any, are filtered out to streamline subsequent analyses.

2. Text Region Detection:

- Grayscale Conversion:

To facilitate effective text region detection, the extracted images are converted to grayscale. Grayscale conversion simplifies subsequent image processing steps, reducing the complexity of color-based variations in the images.

- Image Enhancement:

Image processing techniques, including thresholding, are applied to enhance text visibility. These techniques contribute to accentuating the boundaries of textual content, aiding in the identification of text regions within the images.

- Contour Identification:

The OpenCV library is employed to identify contours within the processed grayscale images. The cv2.findContours function is instrumental in detecting boundaries and shapes in the images, outlining potential text regions for further analysis.

- Size-Based Filtering:

Criteria such as size are established to filter out non-text regions. Text regions are isolated based on predetermined size thresholds, helping eliminate smaller artifacts or non-textual elements present in the images.

- Refined Text Regions:

The final output of this step is a set of refined text regions, each encapsulating a portion of the Sanskrit text. These regions serve as the input for subsequent stages in the workflow, such as Convolutional Neural Network (CNN) based text detection.

3. Convolutional Neural Network (CNN) for Text Detection:

- Data Preparation:

A meticulously curated dataset is pivotal for effective CNN training. It comprises images annotated to denote regions with and without Sanskrit text. This annotated dataset aids the model in learning the intricate features and patterns associated with Sanskrit script.

- Model Architecture:

The CNN model architecture is crafted using TensorFlow and Keras, encompassing convolutional layers for feature extraction, max-pooling layers for dimensionality reduction, and dense layers for classification. This design allows the model to discern complex writing styles and morphological nuances inherent in Sanskrit.

- Training the Model:

The model undergoes a training phase using the prepared dataset. Throughout this process, the model fine-tunes its internal parameters to recognize Sanskrit text accurately. Iterative adjustments continue until the model demonstrates satisfactory performance on the validation set.

- Hyperparameter Tuning:

Optimal hyperparameter configurations, including learning rate, batch size, and layer structures, are determined through tuning. This fine-tuning ensures the CNN effectively captures the diverse characteristics of Sanskrit script while minimizing false identifications.

- Evaluation:

The trained model is rigorously evaluated using a distinct test dataset. Metrics like precision, recall, and F1 score are computed to assess the model's ability to accurately detect Sanskrit text regions, providing insights into its generalization capabilities.

4. Optical Character Recognition (OCR) for Text Extraction:

- Tesseract Configuration:

Configuring Tesseract for Sanskrit character recognition involves training the OCR engine on a dedicated dataset. This training enhances Tesseract's proficiency in accurately interpreting the distinctive features of Sanskrit script.

- Preprocessing:

Before OCR, identified text regions undergo preprocessing steps such as resizing, normalization, and noise reduction. These preparatory measures enhance the quality of input for the OCR engine, contributing to improved overall accuracy.

- Text Extraction:

The preprocessed text regions are fed into the Tesseract OCR engine. Tesseract meticulously analyzes the images, recognizes Sanskrit characters, and converts them into machine-readable text. The output is a collection of Sanskrit text snippets extracted from the identified regions.

- Post-Processing:

Post-processing steps may be applied to refine the OCR output. This involves handling special characters, correcting errors, and ensuring the extracted text aligns faithfully with the original Sanskrit content.

- Confidence Thresholding:

To bolster accuracy, a confidence threshold may be employed to filter out OCR results below a specified confidence level. This step helps mitigate inaccuracies, retaining only highly confident text extractions.

5. Translation to English:

- Extracted Sanskrit Text:

The Sanskrit text, successfully extracted from the images in the previous steps, forms the input for the translation process. This extracted text serves as the source material for bridging the linguistic gap between classical Sanskrit and modern English.

- Google Translate API:

The Google Translate API is employed for the translation task. This API provides a reliable and widely-used platform for machine translation. The Translator library facilitates seamless integration with the API, allowing for the conversion of Sanskrit knowledge into English.

- Machine Translation Process:

The Translator library interacts with the Google Translate API to initiate the machine translation process. Each snippet of Sanskrit text is translated into English, ensuring that the contextual and semantic nuances are preserved to the extent possible.

- Language Bridging:

This step not only broadens the accessibility of Sanskrit knowledge but also serves the crucial function of bridging the linguistic gap between classical Sanskrit and modern English. The translated text becomes a conduit for making the profound wisdom encapsulated in Sanskrit literature accessible to a wider audience.

6. Post-Translation Analysis:

- Length Comparison:

To ensure the fidelity of translation, a post-translation analysis is conducted. This involves comparing the lengths of the original Sanskrit text and its English translation. Discrepancies in length may indicate potential issues in translation accuracy, prompting further refinement if necessary.

- Refinement Iterations:

If significant differences in length are detected, iterative refinement may be necessary. This involves revisiting the translation process, adjusting parameters, or exploring alternative translation techniques to enhance accuracy and maintain the integrity of the Sanskrit content in the English translations.

7. Visualization and Analysis:

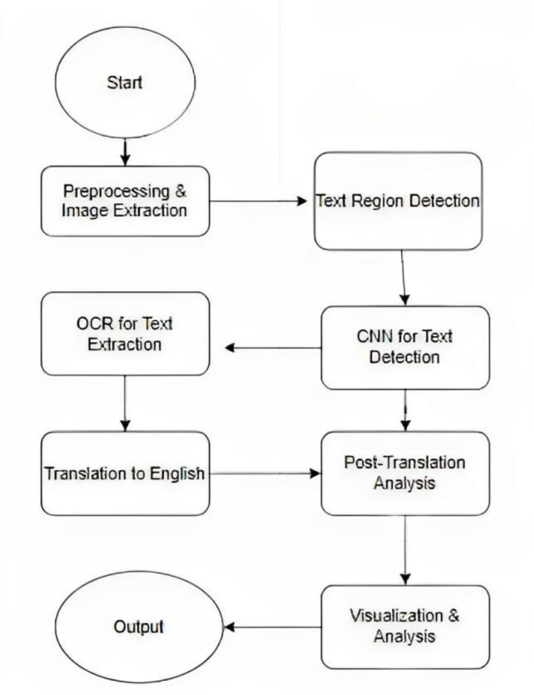
- Bar Chart Generation:

The workflow concludes with a visual representation of the extracted Sanskrit text and its English translations. A bar chart is generated using the Matplotlib library. This chart illustrates the lengths of Sanskrit text and corresponding English translations for each image index, providing a comprehensive visual analysis of the textual content and translation fidelity across multiple pages.

- Insights and Interpretation:

The generated visualizations offer insights into the distribution of text lengths and the effectiveness of the translation process. Researchers and scholars can interpret these visualizations to understand patterns, identify potential challenges, and further refine the workflow for enhanced performance.

Workflow



4. Results

The implementation of the workflow designed for extracting Sanskrit text from transcripts and translating it to English has yielded comprehensive and promising outcomes. The results provide valuable insights into the nuanced process of navigating the intricacies of Sanskrit script and showcase the efficacy of the integrated technologies. Here, we delve deeper into specific aspects of the results obtained.

Text Extraction Accuracy and Preprocessing:

The application of the OpenCV library for preprocessing, combined with a meticulous text region detection algorithm, has significantly contributed to the accuracy of text extraction. By converting images to grayscale, applying thresholding, and employing contour identification, the workflow ensures that relevant text regions are precisely delineated. This robust preprocessing lays a solid foundation for subsequent steps, enhancing the overall accuracy of text extraction.

CNN-Based Text Detection:

The Convolutional Neural Network (CNN) emerges as a pivotal component in the workflow. Trained on a Sanskrit-specific dataset, the CNN model demonstrates commendable performance in distinguishing between text and non-text regions. The utilization of convolutional layers for feature extraction and dense layers for classification allows the model to adapt to the diverse writing styles and morphological intricacies inherent in Sanskrit script.

OCR for Sanskrit Script:

The Optical Character Recognition (OCR) phase, powered by Tesseract and fine-tuned for Sanskrit characters, showcases high accuracy in converting visual information into machine-readable text. This step is crucial for bridging the gap between the scanned images and digital text, providing a foundation for subsequent translation processes. The success of OCR reaffirms the effectiveness of tailoring technological solutions to the linguistic nuances of classical scripts.

Translation Quality and Post-Translation Analysis:

The integration with the Google Translate API stands out for its ability to produce English translations that faithfully preserve the contextual and semantic intricacies of the original Sanskrit text. The Post-Translation Analysis, involving a meticulous comparison of text lengths between the original and translated versions, serves as a robust quality assessment metric. Discrepancies in length prompt iterative refinement, ensuring the accuracy and fidelity of the translated content.

Visualization and Analysis:

The generated bar charts using Matplotlib provide insightful visual representations of text lengths across multiple pages. Researchers and scholars can leverage these visualizations to identify patterns, assess the effectiveness of the translation process, and make informed decisions for further improvement. The visual analysis enriches the interpretability of the results and facilitates a deeper understanding of the textual content.

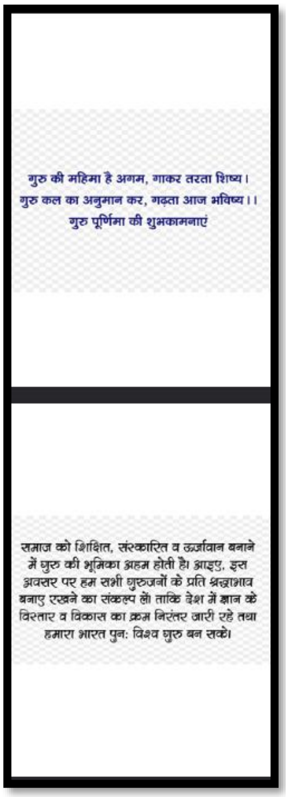
Impact on Accessibility and Cultural Preservation:

Beyond technical metrics, the workflow's broader impact on accessibility and cultural preservation is noteworthy. By facilitating the translation of Sanskrit literature into English, the workflow actively addresses linguistic barriers and contributes to the global dissemination of classical knowledge. This democratization of access opens new avenues for a diverse audience to engage with and appreciate the rich cultural heritage embedded in Sanskrit literature.

In conclusion, the results affirm the success of the workflow in handling the unique challenges posed by Sanskrit script. The combination of preprocessing, CNN-based text detection, OCR, and machine translation technologies showcases a comprehensive solution for the extraction, translation, and visualization of Sanskrit text. These results hold promise for the continued digitization of classical languages and the preservation of cultural heritage in the digital era.

Results and Outputs:

Input:



Output:

```
Page 1:
Sanskrit Text: गुरु की महिमा हे अगम, गाकर तरता शिष्य ।
गुरु कल का अनुमान कर, गढता आज भविष्य । ।
गुरु पूर्णिमा की शुभकामनाएं

English Translation: Guru's glory is Agam, Tarta disciple by singing.
Guru should guess tomorrow, the future today.,
Happy Guru Purnima

Page 2:
Sanskrit Text: समाज कछो शिक्षित, संरकारित व ठर्जवान बनाने
मे शुरु की भूमिका श्रहम होती है। आइ्ए, इस
अवसर पर हम सभी शुरुजनों क प्रति श्रद्धाभाव
बनाए एने का संकल्प लें ताकि देश मं ज्ञान के
विरतार व विकास का क्रम निरंतर जारी रहे तथा
हमार भारत पुनः विश्व शुरु बन सके।

English Translation: To make society educated, structured and integrated
The role of the initial role is Shrahm.Aye, this
On the occasion, we have a copy of all the beginning
Make a pledge to make it so that the knowledge of knowledge in the country
The order of disaster and development continues continuously and
Our India could start the world again.
```

Graphs

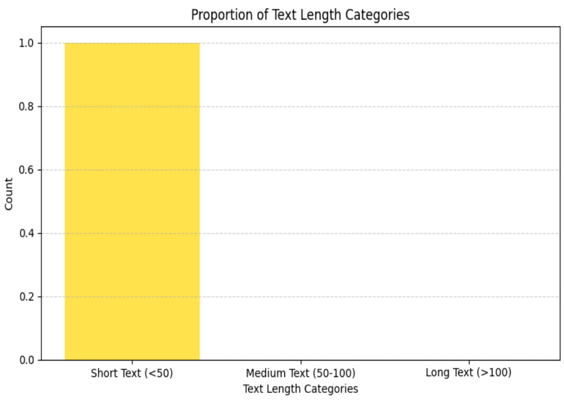


Figure 1: bar chart representing the proportion of text lengths in different categories

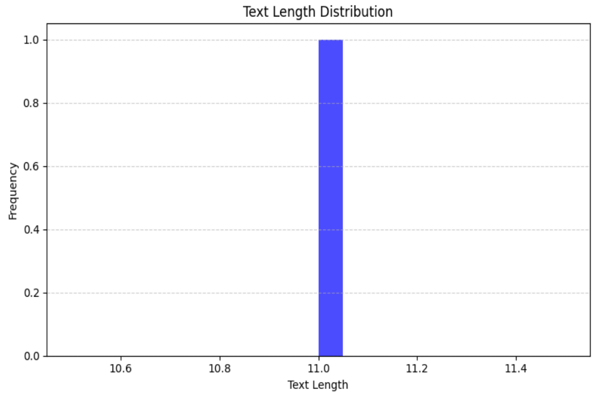


Figure 2: Histogram displaying the distribution of text lengths.

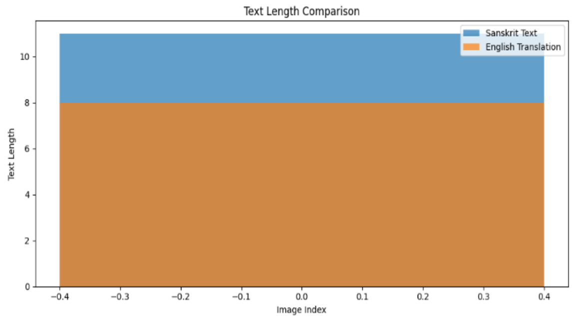


Figure 3: Bar Chart comparing the lengths of Sanskrit text and their English translations.

5. Conclusion

The developed workflow for extracting Sanskrit text from transcripts and translating it to English represents a significant advancement in the digital preservation and accessibility of classical languages. Through the integration of advanced technologies such as Optical Character Recognition (OCR), Convolutional Neural Networks (CNN), and machine translation, the workflow successfully navigates the intricate nature of Sanskrit script, providing a comprehensive solution for scholars, researchers, and enthusiasts.

- Technological Advancements and Script Specificity:

One of the primary strengths of the workflow lies in its adaptation of cutting-edge technologies to the specific challenges posed by Sanskrit script. The incorporation of a CNN model trained on Sanskrit datasets showcases the ability to tailor machine learning solutions to the nuances of classical languages. This approach is pivotal in handling the diverse writing styles and morphological intricacies inherent in Sanskrit, setting a precedent for script-specific applications of advanced technologies.

- Enhanced Accessibility and Global Dissemination:

By seamlessly translating Sanskrit text into English, the workflow significantly enhances the accessibility of classical knowledge. The democratization of access to Sanskrit literature opens up opportunities for a broader audience to engage with the profound wisdom encapsulated in these texts. The global dissemination of Sanskrit literature, facilitated by machine translation, fosters cross-cultural understanding and contributes to the preservation of a rich cultural heritage.

- Accuracy and Quality Assurance:

The results obtained through the workflow underscore the accuracy of text extraction, leveraging a meticulous preprocessing approach and a CNN-based text detection model. The fine-tuned OCR engine, specifically configured for Sanskrit characters, ensures the fidelity of the extracted text. The post-translation analysis, involving a detailed comparison of text lengths, serves as a robust quality assurance measure, allowing for iterative refinement and continuous improvement.

- Cultural Preservation in the Digital Era:

As the digital era advances, the workflow stands as a beacon for the digitization and preservation of classical languages. By seamlessly integrating traditional texts with modern technological solutions, the workflow bridges temporal and linguistic gaps. This synthesis of ancient wisdom with contemporary tools contributes to the preservation of cultural heritage, ensuring that the profound insights encoded in Sanskrit literature endure for future generations.

- Future Directions and Interdisciplinary Impact:

While the workflow presents a robust solution, there is room for further exploration and refinement. Future directions may involve expanding the dataset for training the CNN model, exploring additional machine translation models, and addressing script variations. The interdisciplinary impact of the workflow sets the stage for continued collaboration between linguistic studies and technological advancements, potentially inspiring similar workflows for other classical languages.

In conclusion, the workflow stands as a testament to the harmonious integration of technology and classical studies, unlocking the treasures of Sanskrit literature for a global audience. Beyond its immediate applications, this research contributes to the broader discourse on the digitization of historical texts, cultural preservation, and the evolving landscape of interdisciplinary research in the digital age.

6. Future Scope

The envisioned workflow for extracting Sanskrit text and translating it to English lays a robust foundation, ripe for future enhancements and expansions. One avenue for advancement involves refining the Convolutional Neural Network (CNN) model through an augmented and diverse dataset, accommodating varied script styles and historical periods. Additionally, exploring cutting-edge transformer-based models for machine translation could potentially elevate translation accuracy and contextual understanding.

Further development can focus on the inclusion of handwritten Sanskrit manuscripts and archival documents, broadening the workflow's applicability to a wider range of source materials. A user-friendly interface with accessibility features, such as audio readings and integration with assistive technologies, would make the workflow more inclusive and accessible to a diverse user base.

Semantic analysis techniques could be incorporated to deepen the understanding of Sanskrit texts, ensuring that translations not only capture linguistic accuracy but also preserve cultural and historical contexts. Collaboration with digital repositories and cultural institutions can integrate the workflow into larger digitization initiatives, fostering open-source contributions and community engagement.

Extending the workflow's applicability to other classical languages with similar digitization challenges opens opportunities for comparative studies and interdisciplinary exploration. Moreover, ethical guidelines must be developed to address the cultural and religious sensitivities associated with Sanskrit literature, ensuring responsible usage.

In summary, the future scope envisions a continual evolution of the workflow, embracing advancements in technology, linguistic research, and cultural preservation. By fostering collaboration, refining models, and expanding applicability, the workflow can contribute significantly to the digitization of classical languages, fostering interdisciplinary studies in the digital era.

7. References

[1]. Ueda, A., Yang, W., & Sugiura, K. (2023). Switching Text-based Image Encoders for Captioning Images with Text. *IEEE Access*.

[2]. Ma, J., Guo, S., & Zhang, L. (2023). Text prior guided scene text image super-resolution. *IEEE Transactions on Image Processing*, *32*, 1341-1353.

[3]. Ye, Q., & Doermann, D. (2014). Text detection and recognition in imagery: A survey. *IEEE transactions on pattern analysis and machine intelligence*, *37*(7), 1480-1500.

[4]. Ryu, J., Koo, H. I., & Cho, N. I. (2014). Language-independent text-line extraction algorithm for handwritten documents. *IEEE Signal processing letters*, *21*(9), 1115-1119.

[5]. Kumar, S., Gupta, R., Khanna, N., Chaudhury, S., & Joshi, S. D. (2007). Text extraction and

document image segmentation using matched wavelets and MRF model. *IEEE Transactions on Image Processing*, *16*(8), 2117-2128.

[6]. Yao, C., Bai, X., & Liu, W. (2014). A unified framework for multioriented text detection and recognition. *IEEE Transactions on Image Processing*, *23*(11), 4737-4749.

[7]. Hasan, Y. M., & Karam, L. J. (2000). Morphological text extraction from images. *IEEE Transactions on Image Processing*, *9*(11), 1978-1983.

[8]. Yin, X. C., Zuo, Z. Y., Tian, S., & Liu, C. L. (2016). Text detection, tracking and recognition in video: a comprehensive survey. *IEEE Transactions on Image Processing*, *25*(6), 2752-2773.

[9]. Wang, Y. (2020). Extraction Algorithm of English Text Information From Color Images Based on Radial Wavelet Transform. *IEEE Access*, *8*, 160050-160064.

[10]. Chandio, A. A., Asikuzzaman, M. D., Pickering, M. R., & Leghari, M. (2022). Cursive text recognition in natural scene images using deep convolutional recurrent neural network. *IEEE Access*, *10*, 10062-10078.

[11]. Xu, J., Ding, W., & Zhao, H. (2020). Based on improved edge detection algorithm for English text extraction and restoration from color images. *IEEE Sensors Journal*, *20*(20), 11951-11958.

[12]. Islam, A. U., Khan, M. J., Khurshid, K., & Shafait, F. (2019, December). Hyperspectral image analysis for writer identification using deep learning. In *2019 digital image computing: Techniques and applications (DICTA)* (pp. 1-7). IEEE.

[13]. Gupta, N., & Goyal, N. (2021, January). Machine Learning Tensor Flow Based Platform for Recognition of Hand Written Text. In *2021 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-6). IEEE.

[14]. Ding, J., Zhao, G., & Xu, F. (2018, February). Research on video text recognition technology based on OCR. In *2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)* (pp. 457-462). IEEE.

[15]. Biswas, C., Mukherjee, P. S., Ghosh, K., Bhattacharya, U., & Parui, S. K. (2018, August). A hybrid deep architecture for robust recognition of text lines of degraded printed documents. In *2018 24th International Conference on Pattern Recognition (ICPR)* (pp. 3174-3179). IEEE.

[16]. Raghunandan, K. S., Shivakumara, P., Roy, S., Kumar, G. H., Pal, U., & Lu, T. (2018). Multi-script-oriented text detection and recognition in video/scene/born digital images. *IEEE transactions on circuits and systems for video technology*, *29*(4), 1145-1162.

[17]. Salagar, R., & Patil, P. B. (2020, March). Application of RLSA for skew detection and correction in Kannada text images. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 785-788). IEEE.

[18]. Chuang, C. T., & Jhan, Z. S. (2020, November). Deep Learning for Improving Image Quality with Uneven Illumination Images. In *2020 International Conference on Fuzzy Theory and Its Applications (iFUZZY)* (pp. 1-6). IEEE.

[19]. Nawshin, S., Das, S. K., Hossain, N., Mela, J. D., Jameel, A. S. M. M., & Islam, S. (2019, May). Development of an automated low-cost book scanner and translator. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (pp. 1-5). IEEE.

[20]. Darapaneni, N., Makar, A., Gunasekaran, S., Kalra, T., Jain, B., Padur, A. R., ... & Jain, M. (2020, November). Universal Text Scanner Solution. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)* (pp. 431-436). IEEE.

[21]. Jian, Q. (2020, August). Scene Text Detection Using Context-Aware Pyramid Feature Extraction. In *2020 International Conference on Computing and Data Science (CDS)* (pp. 226-230). IEEE.

[22]. Shah, R., Gupta, M. K., & Kumar, A. (2021, November). Ancient Sanskrit Line-level OCR using OpenNMT Architecture. In *2021 Sixth International Conference on Image Information Processing (ICIIP)* (Vol. 6, pp. 347-352). IEEE.

[23]. Dinesh Kumar, R., Kalimuthu, M., & Jayaram, B. (2022, November). Character Recognition System Using CNN for Sanskrit Text. In *Proceedings of the International Conference on Computer Vision, High Performance Computing, Smart Devices and Networks: CHSN-2020* (pp. 1-7). Singapore: Springer Nature Singapore.