**Chapter 2**

**REVIEW OF RELATED LITERATURE**

* 1. **Introduction**

Stenography, also known as shorthand writing, is most likely the most recognizable action or method of capturing spoken words by writing in shorthand and utilizing a stenotype machine. A manual or written shorthand is a means of writing quickly by replacing letters, sounds, words, or sentences with characters, abbreviations, or symbols. It allows for quick or brief communication and representation. Machine shorthand is another word for writing created by a stenotype, which is an atypical keyboard.

There are several options for stenographers and secretaries, as well as journalists and court reporters, who can take notes quickly with a pen and afterwards convert their notes from shorthand to text. As a result, they claim that shorthand is a useful tool; it is also useful for personal usage. However, due to the enormous number of letters or phrases to be copied, transcription of shorthand writing takes time and might be confusing. Another issue is that only a stenographer is capable of understanding and translating shorthand writing. What if a stenographer is not present to interpret a document? A technology such as optical character recognition (OCR) plays an important part in resolving these issues. This research offers the associated literature of works that make use of a corresponding technology to recognize characters from transcriptions of printed or written texts. In this chapter, we have gathered some related studies and literature on the application of various methodologies that might be useful for a better understanding of the Gregg shorthand translation utilizing OCR.

* 1. **Related Studies**

Optical character recognition (OCR) is the technique of classifying optical patterns included in a digital image that correlate to alphanumeric or other characters that are acquired by optical means, typically a scanner or a camera. Segmentation, feature extraction, and classification are critical phases in character recognition. One article said that OCR has gained increasing attention in both academic research and industry (Chaudhuri et al., 2016). According to Chugh and Arya (2017), OCR is classified into two types: handwritten character recognition and printed character recognition. Due to the variety of human handwriting styles and traditions, the former is more difficult to accomplish than the latter. All recognition approaches are significantly dependent on the type of data to be recognized. To distinguish the characters written by various users, the recognition procedure has to be considerably more rapid and precise. It has been shown that when a template matching approach is applied for the recognition of English character images, recognition improves.

An example of an OCR application is the study of Nour I. Ismail and Elhafiz Mustafa (2022), in which they studied the idea of building a system that recognizes very quickly the common Arabic names. A Probabilistic Neural Network was used to holistically recognize 15 names whose frequency in the data set is 60 or above and represent 41% of the whole data set. By choosing an appropriate rejection threshold, the network was able to recognize 64% of the 20 names where the error was less than 1%.

Another application is the work of Jahandad et al. (2019) on two Convolutional Neural Network Architectures; Inception-v1 and Inception-v3 and proved that these two algorithms can be successfully used to verify individuals in an organization using handwritten signatures and to classify genuine and skilled forged signatures of 1000 users. Each user having 24 genuine and 30 forged signatures. The experimental results from this effort on the Synthetic Signature Database of low-resolution images suggest that Inception-v3 can outperform Inception-v1 on high-resolution 3-dimensional images like ImageNet. According to the creators of Inception-v3, models with higher resolution receptive fields result in dramatically enhanced recognition performance. Inception-v1 with a 22-layer deep network outperformed Inception-v3 with a 42-layer deep network for existing low-resolution input pictures. Their research is vital for effectively identifying persons in an organization, which is critical since companies rely on biometric technologies for individual verification.

Vianny et al. (2022) presented a project that employed CNN and LSTM (BLSTM) for handwriting recognition, and an encoder-decoder employing LSTM for language recognition. They used a combination of handwriting recognition and language translation to accurately recognize handwritten words and translate them into one of India's native languages (Hindi). Their study is significant in the field of handwriting recognition systems for various languages and scripts since it is difficult for everyone to understand every handwriting because each individual has a unique handwriting and even forgeries are not always correct.

A developing multidisciplinary challenge at the junction of computer vision, natural language processing, and artificial intelligence (AI) is generating a natural language description from an image. Image-to-text conversion is one such example. Image-to-text generation is a vital multidisciplinary field that combines computer vision and natural language processing. It also serves as the technological foundation for a number of critical applications. However, in tasks like image classification, the content of an image is often basic, consisting of a prominent object to be categorized. When we want computers to interpret complicated scenarios, the problem might become considerably more difficult. One such task is image captioning (He & Deng, 2017). An example of this is a system that uses optical character recognition to convert a shorthand writing into English text.

Recently, some new study on this topic has been proposed. Kumar et al. (2022) investigated word recognition of handwritten Hindi characters and its application to handwritten forms. Their paper used segmentation-based methodologies to present an end-to-end word detection system for the Hindi language. Their suggested architecture employs an end-to-end technique for recognizing handwritten Hindi words from printed forms and translating them into English. Sentiment analysis is conducted on feedback forms utilizing Random Forest algorithms and NLTK packages such as Porter stemmer and Stop words, yielding an accuracy of 88 percent. This program will assist many individuals in rural regions by making it simple to fill out paperwork without having to worry about the language. They would only need to upload a photograph of the completed form in their local language to the app. This would reduce such people's reliance on others while also making the form-filling procedure easier.

Another example is the study of Perin (2021) in which the researcher developed a mobile Android app that could transliterate main Eskaya characters to their equivalent Latin letters. The supervised machine learning model for transliterating Eskaya characters to Latin, which used k-Nearest Neighbors (k-NN) as the method of choice, performed well, as seen by its high accuracy rate of 89.93 percent. In its present form, the software simply transliterates single Eskaya characters. Future work on the API might train the machine learning model to identify all Eskaya characters, which would be in line with the study's objective of spreading Eskaya literacy. As a result, the program may convert entire Eskaya words into simple Abidiha or more complicated Simplit, perhaps assisting in the preservation of the native language for future generations.

**REFERENCE LIST**

Chaudhuri, A., Mandaviya, K., Badelia, P., & Ghosh, S. K. (2016). Optical Character Recognition Systems. *Optical Character Recognition Systems for Different Languages with Soft Computing*, *352*, 9–41. <https://doi.org/10.1007/978-3-319-50252-6_2>

Chugh, R., & Arya, M. A. (2017). Optical character recognition. *International Journal for Research in Applied Science & Engineering Technology*, *5*(6), 2504–2509. <https://www.ijraset.com/fileserve.php?FID=8536>

He, X., & Deng, L. (2017). Deep learning for Image-to-Text generation: A technical overview. *IEEE Signal Processing Magazine*, *34*(6), 109–116. <https://doi.org/10.1109/msp.2017.2741510>

Jahandad, Sam, S. M., Kamardin, K., Amir Sjarif, N. N., & Mohamed, N. (2019). Offline signature verification using deep learning convolutional neural network (CNN) architectures GoogLeNet inception-v1 and inception-v3. *Procedia Computer Science*, *161*, 475–483. <https://doi.org/10.1016/j.procs.2019.11.147>

Kumar, H., Prasad, A., & Rane, N. (2022). Form scanner & decoder : Conversion of text from any application form and its language translation using OCR. *IJARCCE*, *11*(2), 186–194. <https://doi.org/10.17148/ijarcce.2022.11236>

Nour I. Ismail, M., & Elhafiz Mustafa, M. (2022). Recognition of handwritten arabic names using probabilistic neural networks. *IJARCCE*, *11*(7), 31–35. <https://doi.org/10.17148/ijarcce.2022.11705>

Perin, M.A. (2021). ESKAYAPP. An Eskaya-Latin Script OCR Transliteration E-Learning Android Application using Supervised Machine Learning.

Vianny, M. M., Harshitha, K. C., Keerthana, L., Pavithra, S., & Varshitha, Y. (2022). Handwritten recognition with language translation. *International Journal of Advanced Research in Computer and Communication Engineering*, *11*(5), 943–952. <https://doi.org/10.17148/IJARCCE.2022.115202>