# Formal Project Proposal template

Due 27 June 2023

|  |  |
| --- | --- |
| **Name:** | Bouadi Daryne |
| **Programme:** | MSc Data science |
| **Student ID:** | 23012721 |
| **Project Title:** | *Enhancing Handwritten Character Recognition for Berber Scripts Using Deep Learning Techniques* |
| **Problem you are seeking to address:** | *The lack of efficient optical character recognition (OCR) systems for Berber scripts, especially for Tifinagh and Latin Amazigh characters, is the main issue that this research aims to solve. The underrepresentation of Berber scripts in current OCR systems restricts the digital accessibility and utilization of these languages. This problem limits the use of Berber languages in contemporary technology contexts and hinders their promotion and preservation in digital media. The project's goal is to create and assess a range of deep learning models for character recognition in order to determine the best ways to improve OCR systems that are suited to the particular requirements of these scripts.* |
| **Aim and scope:** | *The aim of this project is to advance the recognition capabilities of OCR systems for Tifinagh and Latin Amazigh scripts by developing and testing a range of deep learning models, including Convolutional Neural Networks (CNNs), Capsule Networks, and Vision Transformers. This research will fill the gap in digital tools that support the recognition of minority languages like Berber, providing more inclusive technological solutions.*  *The project's scope includes using the Berber-MNIST dataset for performance evaluation, model construction, and data preparation. In addition to improving the technical capabilities of currently in use OCR systems, the project will assist Berber scripts' digital presence and accessibility, which will benefit linguistic and cultural studies more broadly. To guarantee that the research is in line with current academic and technological standards, relevant research on deep learning applications, OCR technology, and minority language preservation will be studied.* |
| **Project objectives:** | *1. To develop and implement various deep learning models for recognizing Tifinagh and Latin Amazigh characters.*  *2. To evaluate the performance of each model in terms of accuracy, stability, and suitability for the unique characteristics of the Berber scripts.*  *3. To provide recommendations for enhancing OCR systems tailored to Berber scripts.*  *4. To contribute to the preservation and promotion of Berber languages by improving digital tools for their recognition.* |
| **Expected project outcomes:** | *the expected outcome of this project is a set of optimized deep learning models capable of accurately recognizing handwritten Tifinagh and Latin Amazigh characters. These models will be evaluated based on their accuracy, stability, and adaptability to the unique features of the Berber scripts. The project will also produce a comprehensive analysis of the strengths and weaknesses of each model, along with practical recommendations for developing effective OCR systems tailored to Berber scripts.*  *The developed models and findings will help overcome the digital exclusion of Berber languages by enhancing their representation in digital formats, promoting linguistic diversity, and ensuring that Berber scripts are accessible to speakers, learners, and researchers.* |
| **Brief review of relevant literature:** | *Recent research has made significant advanced in the area of recognizing handwritten characters through the utilization of different machine learning methods, each tailored to specific strengths and uses. Conventional techniques, like Convolutional Neural Networks (CNNs), have been the primary method used in tasks involving the recognition of images. CNNs are highly successful in identifying hand-drawn numbers and letters due to their ability to understand and apply visual patterns from extensive data sets. These networks utilize convolutional filters arranged in layers to identify edges, shapes, and textures in images, which is why they are perfect for visual pattern recognition. LeCun et al. (1998) initially popularized this method for recognizing digits, and later on, Krizhevsky et al. (2012) extended it for tasks on a larger scale such as ImageNet classification. CNNs' standard application in the field is attributed to their adaptability and expandability.Expanding on this base, scientists have designed more advanced CNN structures to address constraints faced by conventional models. An instance would be the ResNet model, which was presented by He et al. (2016), includes residual connections that enable the training of even deeper networks. These connections efficiently avoid specific layers, avoiding problems such as disappearing gradients that may arise when backpropagating in deep networks. ResNet has greatly enhanced image recognition tasks, such as recognizing handwritten characters, by enabling the network to learn more intricate patterns without compromising performance, through maintaining stability in performance and improving accuracy. methods like Capsule Networks have been suggested to tackle particular difficulties in capturing spatial hierarchies and relationships among image features. Introduced by Sabour et al. (2017), Capsule Networks provide a unique method to represent these connections, especially useful for identifying intricate handwritten scripts that rely on the spatial layout of characteristics. Capsule Networks maintain spatial hierarchies, in contrast to traditional CNNs that can lose them with pooling layers, possibly resulting in improved performance.*  *Similarly, Vision Transformers have also become a hopeful replacement for extensive image recognition duties, such as recognizing handwritten characters. Dosovitskiy et al. (2020) created the Vision Transformer model, which utilizes a self-attention mechanism to concentrate on various areas of an image, capturing detailed patterns and connections that conventional models may miss. This approach has demonstrated impressive outcomes in different image recognition assignments, indicating its possible suitability for recognizing handwritten characters, especially in intricate scripts. Further studies have also looked into creating Optical Character Recognition (OCR) systems designed for individual languages and writing systems, in addition to the aforementioned models. An example is the use of hybrid methods that blend various techniques to enhance OCR precision for Arabic writing (Al-Taani & Al-Shatnawi, 2013). These hybrid models capitalize on the advantages of different techniques to address the distinct characteristics of varied scripts, enhancing overall accuracy in recognition. Chiang et al. (2019) further explored this topic by creating techniques that improve OCR precision for intricate handwritten texts, showcasing the flexibility of OCR systems across various language contexts. The importance of regularization methods in boosting model efficiency and generalization has also been highlighted in research. Rifkin and Lippert (2007) emphasized that regularization aids in the prevention of overfitting by penalizing overly intricate models, thus promoting simpler models that perform well on new data. In the realm of handwritten character recognition, data variability is significant, and overfitting is a frequent concern. Recent studies have assessed the machine learning approaches for recognizing handwritten text, emphasizing the significance of adjusting these methods for minority languages and scripts (Sarkar et al., 2022). These reviews indicate an increasing awareness of the importance of inclusive digital tools that cater to a wider variety of languages, especially those with fewer resources available.*  *Furthermore, Tesseract and similar OCR engines have shown the feasibility of incorporating these advanced techniques into practical everyday use scenarios. Tesseract, created by Smith in 2007, has been greatly improved by integrating several machine learning techniques to enhance accuracy and flexibility in recognizing different scripts.  collectively, these studies offer a thorough summary of the progress in recognizing handwritten characters. They demonstrate how different machine learning models, including CNNs and Vision Transformers, can be utilized to improve digital accessibility for underrepresented languages like Berber. These developments help preserve and promote minority languages while also leading to more inclusive technology solutions in a global digital environment.*  ***References:***  *- Al-Taani, A.T. and Al-Shatnawi, A., 2013. Arabic Optical Character Recognition Using Hybrid Approach. \*International Journal of Advanced Computer Science and Applications\*.*  *- Chiang, Y.-J., Huang, J.-H., Hsu, H.-P., and Tsai, S.-S., 2019. Advancing OCR Accuracy for Complex Handwritten Text. \*Pattern Recognition Letters\*, 125, pp.56-64.*  *- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N., 2020. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. \*arXiv preprint\*.*  *- He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep Residual Learning for Image Recognition. In: \*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition\*, Las Vegas, NV, USA, 27-30 June 2016, pp.770-778.*  *- Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks. In: \*Advances in Neural Information Processing Systems\*, 25, pp.1097-1105.*  *- LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-Based Learning Applied to Document Recognition. \*Proceedings of the IEEE\*, 86(11), pp.2278-2324.*  *- Rifkin, R. and Lippert, R.A., 2007. Notes on Regularized Least Squares. \*MIT Technical Report\*.*  *- Sabour, S., Frosst, N. and Hinton, G.E., 2017. Dynamic Routing Between Capsules. In: \*Advances in Neural Information Processing Systems\*, 30, pp.3856-3866.*  *- Sarkar, A., Mandal, B., Ghosh, S., and Gupta, D., 2022. A Comprehensive Survey on Machine Learning Approaches for Handwritten Text Recognition. \*Pattern Recognition\*, 124, pp.108-123.*  *- Smith, R., 2007. An Overview of the Tesseract OCR Engine. In: \*Proceedings of the Ninth International Conference on Document Analysis and Recognition\*, Curitiba, Brazil, 23-26 September 2007, pp.629-633.* |
| **Testing and Evaluation:** | *The project makes use of anonymized, publicly accessible pre-existing datasets rather than gathering new data. i won't gather, store, or handle any personal information. But the effort further ensures that the data is used with respect and recognizes the significance of cultural sensitivity when working with scripts that represent minority languages.* |
| **Ethical considerations:** | *This project does not require gathering original data; it utilizes publicly available and anonymized existing datasets. There will be no collection, storage, or processing of any personal data.. However, the project acknowledges the importance of cultural sensitivity when working with scripts representing minority languages and ensures respectful use of the data.* |
| **Project plan:** |  |
| **Signature:** |  |

Excluding your project title and references, your proposal should be a maximum of 1000 words in length. If using this template, then it would be at most 3 pages excluding references.