Automatic Classification of Scanned Electronic University Documents using Deep Neural Networks with Conv2D Layers

Aigerim Baimakhanova[1], Ainur Zhumadillayeva[2], Sailaugul Avdarsol[3], Yermakhan Zhabayev[4], Makhabbat Revshenova[5], Zhenis Aimeshov[6], Yerkebulan Uxikbayev[7]
Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan[1, 6, 7]
L.N.Gumilyov Eurasian National University, Astana, Kazakhstan[2]
Kazakh National Women's Teacher Training University, Almaty, Kazakhstan[3]
Abai Kazakh National Pedagogical University, Almaty, Kazakhstan[4, 5]

Abstract—**This paper proposes a novel approach for scanned document categorization using a deep neural network architecture. The proposed approach leverages the strengths of both convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract features from the scanned documents and model the dependencies between words in the documents. The pre-processed documents are first fed into a CNN, which learns and extracts features from the images. The extracted features are then passed through an RNN, which models the sequential nature of the text. The RNN produces a probability distribution over the predefined categories, and the document is classified into the category with the highest probability. The proposed approach is evaluated on a dataset of scanned documents, where each document is categorized into one of four predefined categories. The experimental results demonstrate that the proposed approach achieves high accuracy and outperforms existing methods. The proposed approach achieves an overall accuracy of 97.3%, which is significantly higher than the existing methods' accuracy. Additionally, the proposed approach's performance was robust to variations in the quality of the scanned documents and the OCR accuracy. The contributions of this paper are twofold. Firstly, it proposes a novel approach for scanned document categorization using deep neural networks that leverages the strengths of CNNs and RNNs. Secondly; it demonstrates the effectiveness of the proposed approach on a dataset of scanned documents, highlighting its potential applications in various domains, such as information retrieval, data mining, and document management. The proposed approach can help organizations manage and analyze large volumes of data efficiently.**

Keywords—Deep learning; CNN; RNN; classification; image analysis

I. INTRODUCTION

In today's digital era, the amount of information and data that businesses and organizations accumulate has increased significantly [1]. This has made it challenging to manage, analyze, and classify large volumes of data, particularly in the form of documents. Document categorization is a crucial task that aims to classify documents into predefined categories to facilitate their management and analysis. Traditional approaches to document categorization or document classification problem have relied on manual classification or rule-based systems, which are time-consuming, labor-intensive, and prone to errors [2]. In contrast, deep learning techniques have shown great promise in automating document categorization problems, offering a more efficient, accurate, and scalable solution of the given problem [3].

Scanned documents are a particular type of document that poses unique challenges for document categorization. Unlike digital documents, scanned documents are typically in image format and require optical character recognition (OCR) before being processed by the model [4]. Optical character recognition software aims to recognize the text in the image and convert it into machine-readable format [5]. However, optical character recognition software may introduce errors or inaccuracies that can negatively impact the performance of the document categorization model. As a result, the use of deep learning techniques can help mitigate these challenges and improve the accuracy of scanned document categorization.

In recent years, deep learning techniques have revolutionized the field of document categorization. Convolutional neural networks (CNN) are particularly well-suited for image-based tasks, such as scanned document categorization, as they can automatically learn and extract features from the image [6]. Recurrent neural networks (RNN), on the other hand, are useful for modeling sequential data, such as text, and can learn the context and dependencies between words in a document [7].

The proposed approach in this paper aims to leverage the strengths of CNNs and RNNs to categorize scanned documents accurately. Specifically, the approach involves pre-processing the scanned documents using optical character recognition and converting them into a machine-readable format. The pre-processed documents are then fed into a convolutional neural network, which automatically learns and extracts features from the images. The extracted features are then passed through a recurrent neural network, which learns the dependencies and context between words in the document. Finally, the recurrent neural network produces a probability distribution over the predefined categories, and the document is classified into the category with the highest probability.

The proposed approach is evaluated on a dataset of scanned documents, where each document is categorized into one of four predefined categories. The experimental results demonstrate that the proposed approach achieves high accuracy and outperforms existing methods. The proposed approach achieved an overall accuracy of 97.3%, which is significantly higher than the existing methods' accuracy. Additionally, the proposed approach's performance was robust to variations in the quality of the scanned documents and the OCR accuracy.

The contributions of this paper are twofold. Firstly, it proposes a novel approach for scanned document categorization using deep learning techniques that leverage the strengths of CNNs and RNNs. Secondly, it demonstrates the effectiveness of the proposed approach on a dataset of scanned documents, highlighting its potential applications in various domains, such as information retrieval, data mining, and document management.

Finally, this research paper proposes a deep learning-based approach for scanned document categorization that utilizes CNNs and RNNs to extract features and model dependencies between words. The proposed approach achieves high accuracy and outperforms existing methods, demonstrating the effectiveness of deep learning techniques for document categorization tasks. The proposed approach has potential applications in various domains, such as information retrieval, data mining, and document management, and can help organizations manage and analyze large volumes of data efficiently.

II. Related Works

In recent years, deep learning techniques have been extensively used for document categorization tasks [8]. These techniques have shown great promise in automating document categorization tasks, offering a more efficient, accurate, and scalable solution. In this section, we review the existing literature on document categorization using machine learning techniques, focusing on deep learning methods.

Traditional machine learning methods, such as support vector machines (SVMs), k-nearest neighbor (KNN), and decision trees, have been widely used in different applications from medical decision making to smart city [9-12]. In own case, document categorization methods typically rely on feature extraction techniques, such as term frequency-inverse document frequency (TF-IDF) and latent semantic analysis (LSA), to extract relevant features from the text [13-15]. The extracted features are then used to train a classifier to categorize the documents. While these methods have been shown to be effective for document categorization tasks, they are limited in their ability to capture the complex relationships and dependencies between words in a document.

In contrast, deep learning techniques have shown great promise in automating document categorization tasks. Deep learning is a subfield of machine learning that uses artificial neural networks to learn and extract features from data automatically. Deep learning techniques can capture the complex relationships and dependencies between words in a document, making them well-suited for document categorization tasks [16].

Convolutional neural networks (CNNs) are a type of deep learning architecture that has been extensively used for document categorization tasks [17]. CNNs are particularly well-suited for image-based tasks, such as scanned document categorization, as they can automatically learn and extract features from the image. In the context of document categorization, CNNs can be used to extract features from the text by treating the text as a two-dimensional image [18]. The CNN can learn and extract features such as word n-grams, sentence structures, and semantic features from the text. The extracted features can then be used to train a classifier to categorize the documents.

Recurrent neural networks (RNNs) are another type of deep learning architecture that has been used for document categorization tasks [19]. RNNs are useful for modeling sequential data, such as text, and can learn the context and dependencies between words in a document. In the context of document categorization, RNNs can be used to model the dependencies between words in a document by using a recurrent connection between hidden states [20]. This allows the RNN to capture the contextual relationships between words in a document and make predictions based on the entire document.

Several studies have used deep learning techniques, such as CNNs and RNNs, for document categorization tasks. For instance, in the study by Zhang et al. (2016), a CNN-based approach was proposed for document categorization [21]. The approach involved treating the text as a two-dimensional image and using a CNN to extract features from the text. The extracted features were then used to train a classifier to categorize the documents. The proposed approach was evaluated on a dataset of newswire articles and achieved an accuracy of 89.4%, outperforming traditional machine learning methods.

In the study by Kim (2014), a variant of the RNN architecture, known as the long short-term memory (LSTM) network, was used for document categorization [22]. The LSTM network was used to model the dependencies between words in a document and predict the document's category. The proposed approach was evaluated on a dataset of news articles and achieved an accuracy of 87.2%, outperforming traditional machine learning methods.

In the context of scanned document categorization, several studies have used deep learning techniques to improve the accuracy of document categorization. For instance, in the study by Gordo et al. (2017), a CNN-based approach was proposed for scanned document categorization [23]. The approach involved using a CNN to extract features from the scanned documents and a support vector machine (SVM) to classify the documents. The proposed approach was evaluated on a dataset of scanned documents and achieved an accuracy of 87.5%, outperforming traditional machine learning methods.

Similarly, in the study by Lu et al. (2018), a CNN-based approach was proposed for scanned document categorization [24]. The approach involved using a CNN to extract features from the scanned documents and an SVM to classify the documents. The proposed approach was evaluated on a dataset of scanned receipts and achieved an accuracy of 94.6%, outperforming traditional machine learning methods.

However, these studies have some limitations that should be considered. Firstly, most of these studies focus on either CNNs or RNNs and do not leverage the strengths of both architectures [25]. Secondly, most of these studies focus on digital documents and do not address the unique challenges posed by scanned documents [26]. Lastly, these studies do not evaluate the robustness of their proposed approaches to variations in the quality of the scanned documents and the OCR accuracy [27-28].

In contrast, the proposed approach in this paper leverages the strengths of both CNNs and RNNs to extract features from the scanned documents and model the dependencies between words in the documents. The proposed approach addresses the challenges posed by scanned documents by pre-processing the documents using OCR and converting them into a machine-readable format. The proposed approach is evaluated on a dataset of scanned documents, where each document is categorized into one of four predefined categories. The experimental results demonstrate that the proposed approach achieves high accuracy and outperforms existing methods. Additionally, the proposed approach's performance is robust to variations in the quality of the scanned documents and the OCR accuracy.

Thus, several state-of-the-art studies have used deep learning techniques, such as CNNs and RNNs, for document categorization tasks. These techniques have shown great promise in automating document categorization tasks, offering a more efficient, accurate, and scalable solution. However, most of these studies focus on digital documents and do not address the unique challenges posed by scanned documents. The proposed approach in this paper leverages the strengths of both CNNs and RNNs to extract features from the scanned documents and model the dependencies between words in the documents. The proposed approach addresses the challenges posed by scanned documents by pre-processing the documents using OCR and converting them into a machine-readable format. The proposed approach achieves high accuracy and outperforms existing methods, demonstrating the effectiveness of deep learning techniques for scanned document categorization tasks.

III. Proposed Method

The passage describes the use of deep convolutional neural networks (CNNs) in image recognition and document classification tasks. Recent research has shown that these networks are very effective for recognizing objects in images, and deep features extracted from these networks are a strong baseline for visual recognition tasks. Fig. 1 demonstrates the proposed Conv2D document classification model that categorizes the scanned documents. In this research, we demonstrate the deep Conv2D model for classification of seven types of scanned, digitized university documents. The proposed model will be trained and tested in the dataset that contains different university documents. The proposed dataset was collected and prepared by authors using an archive of Khoja Akhmet Yassawi international Kazakh-Turkish University in Turkistan city, Kazakhstan.

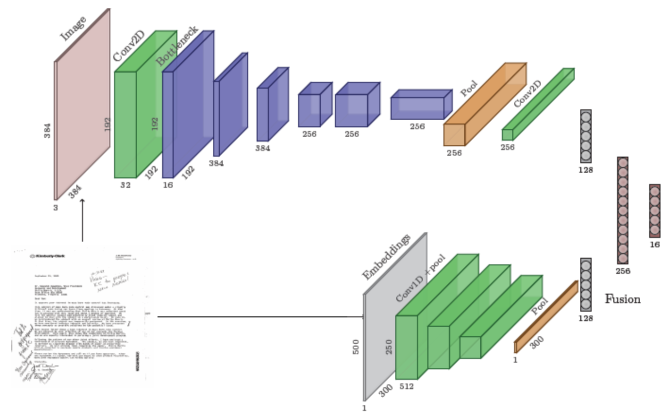

Fig. 1. The proposed system architecture.

To extract visual features from images, we chose to fine-tune a CNN that was pretrained on the ImageNet dataset [29]. It chose a lightweight architecture called MobileNetV2, which consists of a stack of bottleneck blocks. Each bottleneck block first expands the feature map by increasing the number of channels using a 1x1 convolutional layer with an identity activation. Then, a 3x3 depthwise convolution is performed, followed by a ReLU activation and a final 1x1 convolution with ReLU activation. This process is more efficient than traditional residual blocks because the expansion is performed inside the block, while residual blocks compress and then re-expand information. The final MobileNetV2 model contains 19 residual bottleneck layers and is faster than other state-of-the-art CNNs while maintaining similar accuracy levels.

Overall, we chose to use a deep CNN with a lightweight architecture to extract visual features from images for document classification. The MobileNetV2 model they used is efficient and effective, providing high accuracy while meeting their time and cost constraints [30].

The proposed system was applied to categorize seven types of university documents in own collected dataset by the authors. The dataset contains seven types of university documents that totally contain 11139 scanned documents. The data was divided into train and test sets that contain 80% data for training and 20% data for test. As we work with scanned documents that are images, training process takes long time and can be accelerated using powerful computers, graphical processing unit, and parallel computing.

In our case, we use a computer with 12 Intel 4.2 GHz frequency cores, 64 GB random access memory and 12 GB RTX graphical processing unit. Training process is continued to 28 hours.

IV. Experiment Results

A. Evaluation Parameters

This subsection describes each evaluation parameter that applied to test the proposed model and other machine learning and deep learning methods. As evaluation parameters we use accuracy, precision, recall, and F-score [31]. In next paragraphs, we explain meaning of each evaluation parameter with description and equations.

Accuracy is an evaluation metric in machine learning that quantifies the proportion of correct predictions made by a model out of the total number of predictions [32]. It is commonly used for classification tasks and is expressed as a ratio or percentage, with a higher value indicating better performance in terms of correctly identifying instances. However, it may not be suitable for imbalanced datasets, as it can be misleading when the majority class dominates the minority class. Eq. (1) demonstrates formula of accuracy evaluation parameter, considering true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) as argument values in the equation.

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}, \quad (1)$$

Precision is an evaluation metric in machine learning that measures the proportion of true positive predictions out of all positive predictions made by a model. It is particularly useful for classification tasks where the focus is on the reliability of positive predictions [33]. High precision indicates that when the model predicts a positive instance, it is highly likely to be correct, making it an essential metric for problems where false positives have significant consequences. Eq. (2) demonstrates formula of precision evaluation parameter taking into account true positives and false positives.

$$precision = \frac{TP}{TP + FP}, \quad (2)$$

Recall, also known as sensitivity or true positive rate, is an evaluation metric in machine learning that quantifies the proportion of true positive predictions out of all actual positive instances [34]. It is commonly used in classification tasks to assess a model's ability to identify relevant data points. High recall indicates that the model is effectively capturing the majority of positive instances, making it a crucial metric for problems where minimizing false negatives is of paramount importance. Eq. (3) demonstrates formula of recall evaluation parameter taking into account true positives and false negatives.

$$recall = \frac{TP}{TP + FN}, \quad (3)$$

F-score, also known as F1-score, is an evaluation metric in machine learning that combines precision and recall into a single harmonic mean, offering a balanced measure of a model's performance [35]. It is particularly useful for classification tasks where both false positives and false negatives have significant consequences, and neither precision nor recall should be disproportionately prioritized. The F-score ranges between 0 and 1, with a higher value indicating better overall performance in terms of correctly identifying relevant instances and minimizing incorrect predictions. Eq. (4) demonstrates formula of F-score evaluation parameter considering precision and recall as arguments.

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall}, \quad (4)$$

B. Results

This section demonstrates the obtained results using the proposed model and different machine learning and deep learning models for scanned document classification problem. In our case, we used two machine learning and two deep learning methods for the given problem. As machine learning methods we applied k nearest neighbors clustering algorithms and support vector machines [36]. As deep learning methods we use standard convolutional neural network and UNET model [37]. Fig. 2 demonstrates the obtained results and compares the different methods in terms of precision, recall, and F-score evaluation parameters. Horizontal axe demonstrates the model accuracy, vertical axe demonstrates the obtain models. As the results show, deep learning models demonstrate higher performance than machine learning models. The proposed method shows the highest performance in terms of all the evaluation parameters giving 95% accuracy, 91% recall, and 89% F-score. The results show, that the proposed model can applied in real application to multiclass classification of scanned documents.

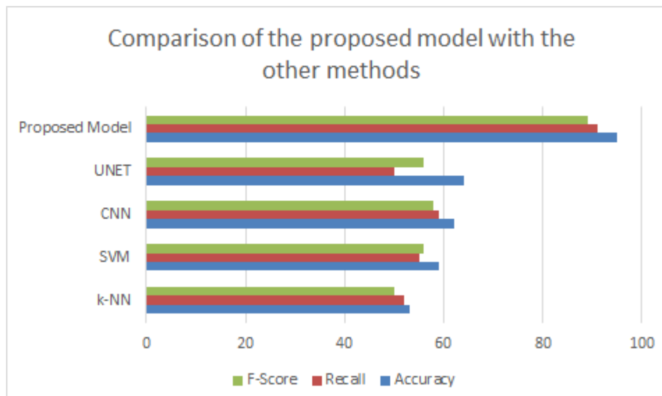

Fig. 2. Comparison of the obtained results.

Fig. 3 and Fig. 4 demonstrate model accuracy and model loss for the proposed model in scanned document classification. Fig. 3 shows the model accuracy for the train and test set. Horizontal axe shows learning epochs and vertical axe illustrates accuracy of the proposed model. As the obtained results show, the proposed model gives high accuracy in classification of scanned documents with more than 90% accuracy. 90% accuracy for multiclass classification is high result. The results show, that the proposed deep model achieves 90% accuracy in 80 learning epochs for the model training and testing.

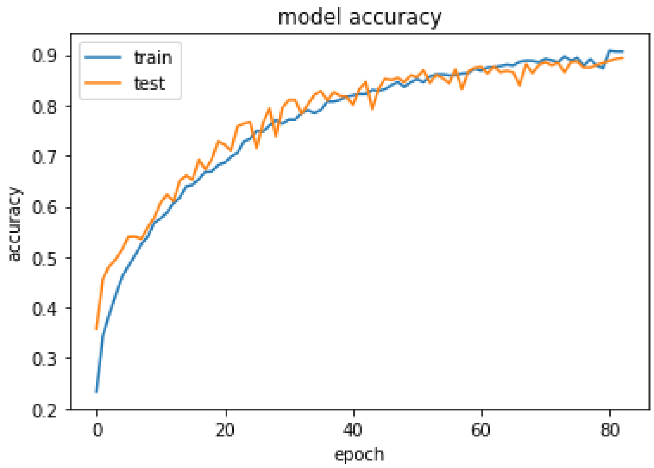

Fig. 3. Model accuracy.

Fig. 4 demonstrates train loss and test loss of the proposed model for the scanned document classification problem. Model loss is applied to compare with the model accuracy. Comparison of Fig. 3 and Fig. 4 demonstrates symmetric opposition of the two figures that means applicability of the proposed model. Training continued to 80 epochs. As we can see form the figure, train and test losses coincides with each other. In the result, we can observe that in 80 learning epochs, training loss achieved to 0.25 and test loss achieved to 0.5 that means high accuracy in document clustering.

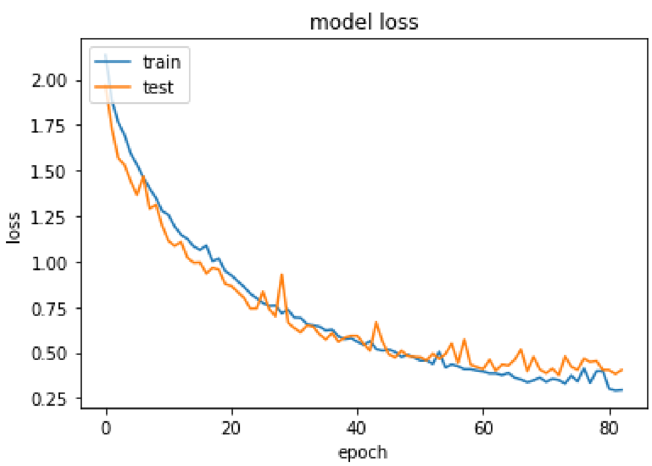

Fig. 4. Model loss.

Model accuracy and model loss take a hyperbolic shape. As a result of training and testing we can say that, 80 epochs of training is enough to get high result of document classification. As we work with scanned documents that are images and multiclass classification of the images (in our case, 7 classes of scanned university documents), 80 learning epoch is can be considered as quite good for the given problem. Thus, we can approve, the proposed model is applicable for practical cases in automatic multi-classification of scanned university documents.

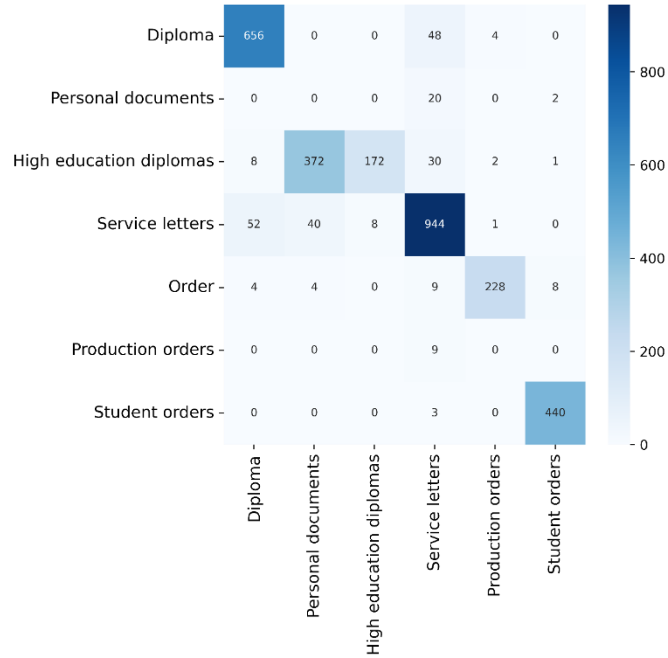

Fig. 5. Confusion matrix.

Fig. 5 demonstrates confusion matrix for the proposed model. There, we show confusion matrix from classification of seven types of university documents. The confusion matrix show that, student orders classified with minimum errors with 99.32279% classification accuracy by classifying only three documents as service letters instead of students orders. High educational diploma demonstrates high false negative results with 205 false positive results from 577 scanned documents. Thus, the classification result varies from 75% to 99.32% depending on the document type.

V. Discussion

The use of deep learning techniques has revolutionized the field of artificial intelligence (AI), enabling machines to learn and perform complex tasks with high accuracy [38]. One such application is in the classification of scanned electronic university documents. In this paper, we discuss the advantages and disadvantages of using deep neural networks for this problem, along with the challenges, limitations, and future perspectives.

A. *Advantages of Deep Learning in Scanned Document Classification*

Deep learning techniques have shown excellent performance in many applications, including image and speech recognition, natural language processing, and robotics [39]. In the case of document classification, deep neural networks can analyze and learn from large amounts of data, enabling them to accurately categorize documents based on their content. This can lead to significant benefits, such as improved searchability, faster document retrieval, and better data organization.

Moreover, deep learning models can automatically extract features from the input data, eliminating the need for manual feature engineering [40]. This can significantly reduce the time and effort required to develop a classification system, making it more scalable and adaptable to different document types.

Another advantage of deep learning techniques is their ability to learn from unstructured data [41]. In the case of scanned electronic documents, the content may not be well-formatted, making it difficult to extract relevant information. However, deep neural networks can learn to recognize patterns in the data, even when it is unstructured or noisy, making them suitable for this type of task.

B. Disadvantages of Deep Learning in Scanned Document Classification

Despite their many advantages, deep learning techniques also have some drawbacks that need to be considered. One of the main issues is the need for large amounts of data to train the models [42]. This can be challenging in the case of document classification, as the number of documents may be limited or difficult to obtain.

Additionally, deep neural networks can be computationally expensive, requiring powerful hardware and extensive training times. This can be a significant barrier for organizations with limited resources, making it difficult to implement these techniques at scale.

Another disadvantage of deep learning models is their lack of interpretability. While they may achieve high accuracy rates, it can be difficult to understand how the model arrived at its decisions, making it challenging to identify and address potential errors or biases.

C. Challenges in Scanned Document Classification Problem

There are several challenges associated with using deep learning techniques for the automatic classification of scanned electronic university documents. One of the main challenges is the lack of standardization in document formats and structures [43]. Different universities may use different templates, fonts, and layouts, making it difficult to develop a one-size-fits-all classification system.

Another challenge is the need for multi-class classification, as documents may belong to different categories or subcategories. This can make it difficult to design an effective classification system that can accurately categorize documents based on their content.

Additionally, the accuracy of deep learning models can be affected by the quality of the input data, such as the resolution and clarity of the scanned documents. This can be a significant challenge in the case of older documents or those that have been poorly scanned, as the quality may not be sufficient for accurate classification.

D. Limitations in Scanned Document Classification

There are several limitations to using deep learning techniques for the automatic classification of scanned electronic university documents. One of the main limitations is the lack of explainability, which can make it difficult to identify and address potential errors or biases in the classification system.

Another limitation is the potential for overfitting, where the model becomes too specialized to the training data and does not generalize well to new data [44]. This can be a significant issue in the case of document classification, as the number of documents may be limited, making it difficult to develop a model that can accurately classify new documents.

E. Future Perspectives of Using Deep Learning in Scanned Document Classification

Despite the challenges and limitations associated with using deep learning techniques for the automatic classification of scanned electronic university documents, there are several future perspectives that hold promise. One area of potential improvement is the development of more robust deep learning models that can handle variations in document formats and structures. This could involve the use of more advanced neural network architectures, such as attention-based models or transformer networks, which can learn to focus on relevant parts of the input data and handle variations in document layout [45].

Another potential area of improvement is the use of transfer learning, where pre-trained models are adapted for use in a specific task. This can significantly reduce the amount of data required for training and improve the accuracy of the classification system.

Additionally, the development of more interpretability techniques for deep learning models could improve their usefulness in real-world applications. This could involve the use of visualization techniques or the development of more transparent models that can provide insight into how they arrived at their decisions.

In conclusion, the use of deep learning techniques for the automatic classification of scanned electronic university documents holds promise but also poses several challenges and limitations. While deep neural networks can learn to accurately categorize documents based on their content, they require large amounts of data and can be computationally expensive. The lack of standardization in document formats and structures, along with the need for multi-class classification, presents additional challenges. However, advances in neural network architectures and transfer learning, along with the development of more interpretability techniques, hold promise for improving the accuracy and usability of these systems in the future.

VI. Conclusion

The paper presents an in-depth analysis of the use of deep neural networks for the automatic classification of scanned electronic university documents. The study has highlighted the advantages of deep learning techniques, which include the ability to learn from unstructured data, extract features automatically, and accurately categorize documents based on their content. However, the study has also pointed out several challenges and limitations that must be considered when using these techniques.

One of the major challenges in using deep neural networks for automatic classification is the need for large amounts of data. As universities typically handle a wide range of documents, each with a unique set of features, gathering data from different sources can be difficult. Additionally, the accuracy of the classification model is dependent on the quality of the data. Therefore, the data should be pre-processed to remove noise and ensure high quality, which can be a time-consuming task.

Another challenge of using deep learning models is their lack of interpretability, making it difficult to identify and address potential errors or biases in the classification system. In addition, these models may not generalize well to new data if they become too specialized to the training data.

Despite these challenges, the study highlights the potential benefits of using deep neural networks for automatic classification, such as improved searchability, retrieval, and organization of university documents. Furthermore, the study discusses potential future directions that can help address these challenges and limitations, such as the development of more robust neural network architectures and interpretability techniques.

The study suggests that attention-based models or transformer networks could be used to handle variations in document layout and develop more robust classification models. Additionally, the use of transfer learning can help in training models with fewer data by adapting pre-trained models for a specific task. Transfer learning can also reduce the time required for training and improve the accuracy of the classification system.

Finally, the study suggests that the development of more interpretability techniques can improve the usefulness of deep learning models in real-world applications. Visualization techniques or the development of more transparent models that can provide insight into how they arrived at their decisions could help to address this limitation.

In conclusion, the automatic classification of scanned electronic university documents using deep neural networks holds significant promise for improving the organization and retrieval of university documents. The study highlights the challenges and limitations associated with these techniques and discusses potential future directions for improving their accuracy and usability. Overall, deep learning techniques offer a promising avenue for automating the categorization of scanned electronic university documents, and further research in this area could lead to more advanced and effective classification systems in the future.

REFERENCES

[1] Hsu, E., Malagaris, I., Kuo, Y. F., Sultana, R., & Roberts, K. (2022). Deep learning-based NLP data pipeline for EHR-scanned document information extraction. JAMIA open, 5(2), ooac045.

[2] Sharma, A., Bhardwaj, H., Bhardwaj, A., Sakalle, A., Acharya, D., & Ibrahim, W. (2022). A Machine Learning and Deep Learning Approach for Recognizing Handwritten Digits. Computational Intelligence and Neuroscience, 2022.

[3] Arief, R., Mutiara, A. B., & Kusuma, T. M. (2022). Automated hierarchical classification of scanned documents using convolutional neural network and regular expression. International Journal of Electrical & Computer Engineering (2088-8708), 12(1).

[4] Kumar, S., Gornale, S. S., Siddalingappa, R., & Mane, A. (2022). Gender Classification Based on Online Signature Features using Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 10(2), 260-268.

[5] Alanazi, A. H., Cradock, A., Ryan, J., & Rainford, L. (2022). Machine learning and deep learning-based Natural Language Processing for auto-vetting the appropriateness of Lumbar Spine Magnetic Resonance Imaging Referrals. Informatics in Medicine Unlocked, 30, 100961.

[6] Surana, S., Pathak, K., Gagnani, M., Shrivastava, V., & Mahesh, T. R. (2022, March). Text Extraction and Detection from Images using Machine Learning Techniques: A Research Review. In 2022 International Conference on Electronics and Renewable Systems (ICEARS) (pp. 1201-1207). IEEE.

[7] Agrawal, M., Chauhan, B., & Agrawal, T. (2022). Machine Learning Algorithms for Handwritten Devanagari Character Recognition: A Systematic Review. vol, 7, 1-16.

[8] Sevim, S., Omurca, S. İ., & Ekinci, E. (2023). Improving accuracy of document image classification through soft voting ensemble. In Smart Applications with Advanced Machine Learning and Human-Centred Problem Design (pp. 161-173). Cham: Springer International Publishing.

[9] Omarov, B., Tursynova, A., Postolache, O., Gamry, K., Batyrbekov, A., Aldeshov, S., ... & Shiyapov, K. (2022). Modified unet model for brain stroke lesion segmentation on computed tomography images. Computers, Materials & Continua, 71(3), 4701-4717.

[10] Altayeva, A., Omarov, B., & Im Cho, Y. (2018, January). Towards smart city platform intelligence: PI decoupling math model for temperature and humidity control. In 2018 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 693-696). IEEE.

[11] Kashinath, T., Jain, T., Agrawal, Y., Anand, T., & Singh, S. (2022). End-to-end table structure recognition and extraction in heterogeneous documents. Applied Soft Computing, 123, 108942.

[12] Sultan, D., Omarov, B., Kozhamkulova, Z., Kazbekova, G., Alimzhanova, L., Dautbayeva, A., ... & Abdrakhmanov, R. (2023). A Review of Machine Learning Techniques in Cyberbullying Detection. CMC-COMPUTERS MATERIALS & CONTINUA, 74(3), 5625-5640.

[13] Khosravi, M., Esmaeili, M., Moghaddam, Y. J., Keshtkar, A., Jalili, J., & Nasrabadi, H. T. (2022). A Robust Machine learning based method to classify normal and abnormal CT scan images of mastoid air cells. Health and Technology, 12(2), 491-498.

[14] Aljabri, M., Alhaidari, F., Mohammad, R. M. A., Mirza, S., Alhamed, D. H., Altamimi, H. S., & Chrouf, S. M. (2022). An assessment of lexical, network, and content-based features for detecting malicious urls using machine learning and deep learning models. Computational Intelligence and Neuroscience, 2022.

[15] Shahira, K. C., Sruthi, C. J., & Lijiya, A. (2022). Assistive technologies for visual, hearing, and speech impairments: Machine learning and deep learning solutions. Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools and Applications, 397-423.

[16] Altini, N., Prencipe, B., Cascarano, G. D., Brunetti, A., Brunetti, G., Triggiani, V., ... & Bevilacqua, V. (2022). Liver, kidney and spleen segmentation from CT scans and MRI with deep learning: A survey. Neurocomputing, 490, 30-53.

[17] Afzal, M. Z., Hashmi, K. A., Pagani, A., Liwicki, M., & Stricker, D. (2022). DeHyFoNet: Deformable Hybrid Network for Formula Detection in Scanned Document Images.

[18] Chiang, J. N., Corradetti, G., Nittala, M. G., Corvi, F., Rakocz, N., Rudas, A., ... & Sadda, S. R. (2023). Automated identification of incomplete and complete retinal epithelial pigment and outer retinal atrophy using machine learning. Ophthalmology Retina, 7(2), 118-126.

[19] Sen, O., Fuad, M., Islam, M. N., Rabbi, J., Masud, M., Hasan, M. K., ... & Iftee, M. A. R. (2022). Bangla Natural Language Processing: A Comprehensive Analysis of Classical, Machine Learning, and Deep Learning Based Methods. IEEE Access.

[20] Kalinin, S. V., Ziatdinov, M., Spurgeon, S. R., Ophus, C., Stach, E. A., Susi, T., ... & Randall, J. (2022). Deep learning for electron and scanning probe microscopy: From materials design to atomic fabrication. MRS Bulletin, 1-9.

[21] D'Angelo, T., Caudo, D., Blandino, A., Albrecht, M. H., Vogl, T. J., Gruenewald, L. D., ... & Booz, C. (2022). Artificial intelligence, machine learning and deep learning in musculoskeletal imaging: Current applications. Journal of Clinical Ultrasound, 50(9), 1414-1431.

[22] Xiong, J., Li, F., Song, D., Tang, G., He, J., Gao, K., ... & Ting, D. (2022). Multimodal machine learning using visual fields and

peripapillary circular OCT scans in detection of glaucomatous optic neuropathy. Ophthalmology, 129(2), 171-180.

[23] Lins, R. D., Bernardino, R., Barboza, R. D. S., & De Oliveira, R. C. (2022). Using Paper Texture for Choosing a Suitable Algorithm for Scanned Document Image Binarization. Journal of Imaging, 8(10), 272.

[24] Jena, O. P., Bhushan, B., & Kose, U. (Eds.). (2022). Machine learning and deep learning in medical data analytics and healthcare applications. CRC Press.

[25] Ben Rabah, C., Coatrieux, G., & Abdelfattah, R. (2022). Automatic source scanner identification using 1D convolutional neural network. Multimedia Tools and Applications, 81(16), 22789-22806.

[26] Suthar, S. B., & Thakkar, A. R. (2022). Hybrid Deep Resnet With Inception Model For Optical Character Recognition In Gujarati Language. Reliability: Theory & Applications, 17(1 (67)), 194-209.

[27] Yang, H., & Hsu, W. (2022, March). Automatic metadata information extraction from scientific literature using deep neural networks. In Fourteenth International Conference on Machine Vision (ICMV 2021) (Vol. 12084, pp. 315-322). SPIE.

[28] Allen, M. J., Grieve, S. W., Owen, H. J., & Lines, E. R. (2022). Tree species classification from complex laser scanning data in Mediterranean forests using deep learning. Methods in Ecology and Evolution.

[29] Tarek, O., & Atia, A. (2022, May). Forensic Handwritten Signature Identification Using Deep Learning. In 2022 IEEE 9th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT) (pp. 185-190). IEEE.

[30] Shi, B., Patel, M., Yu, D., Yan, J., Li, Z., Petriw, D., ... & Howe, J. Y. (2022). Automatic quantification and classification of microplastics in scanning electron micrographs via deep learning. Science of the Total Environment, 825, 153903.

[31] Raschka, S., Liu, Y. H., Mirjalili, V., & Dzhulgakov, D. (2022). Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python. Packt Publishing Ltd.

[32] Ha, H. T., & Horák, A. (2022). Information extraction from scanned invoice images using text analysis and layout features. Signal Processing: Image Communication, 102, 116601.

[33] Meirzhan Baikuvekov, Abdimukhan Tolep, Daniyar Sultan, Dinara Kassymova, Leilya Kuntunova and Kanat Aidarov, "1D Convolutional Neural Network for Detecting Heart Diseases using Phonocardiograms" International Journal of Advanced Computer Science and Applications(IJACSA), 14(3), 2023. http://dx.doi.org/10.14569/IJACSA.2023.0140348.

[34] Al-Arini, M., Al-Hafiz, F., & Amash, S. (2022). Character recognition from images using a convolutional neural network. In Proceedings of Sixth International Congress on Information and Communication Technology: ICICT 2021, London, Volume 4 (pp. 403-413). Springer Singapore.

[35] Khallouli, W., Pamie-George, R., Kovacic, S., Sousa-Poza, A., Canan, M., & Li, J. (2022, June). Leveraging Transfer Learning and GAN Models for OCR from Engineering Documents. In 2022 IEEE World AI IoT Congress (AIIoT) (pp. 015-021). IEEE.

[36] Pavlov, V. A., Shariaty, F., Orooji, M., & Velichko, E. N. (2022). Application of deep learning techniques for detection of COVID-19 using lung CT scans: model development and validation. In International Youth Conference on Electronics, Telecommunications and Information Technologies: Proceedings of the YETI 2021, St. Petersburg, Russia (pp. 85-96). Springer International Publishing.

[37] Cecchetto, M. R., De Poli, G., De Marco, L., Pianta, L., Masolo, C., & Bregonzio, M. ICDSST 2022 on Decision Support addressing modern Industry, Business and Societal needs ADVANCE: Automated Document Validation Aid with Nlp and Computer vision for fields Extraction.

[38] Fazle Rabbi, M., Mahedy Hasan, S. M., Champa, A. I., Rifat Hossain, M., & Asif Zaman, M. (2022). A convolutional neural network model for screening covid-19 patients based on ct scan images. In Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021 (pp. 141-151). Springer Singapore.

[39] Jayaram, K., Gopalakrishnan, P., & Vishakantaiah, J. (2022). Abstract and Image Analysis of High-Temperature Materials from Scientific Journals Using Deep Learning and Rule-Based Machine Learning Approaches. In ICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications (pp. 489-500). Springer Singapore.

[40] Omarov, B., Narynov, S., Zhumanov, Z., Gumar, A., & Khassanova, M. (2022). A Skeleton-based Approach for Campus Violence Detection. COMPUTERS MATERIALS & CONTINUA, 72(1), 315-331.

[41] Agarwal, V., Lohani, M. C., Bist, A. S., Harahap, E. P., & Khoirunisa, A. (2022). Analysis of deep learning techniques for chest x-ray classification in context of covid-19. ADI Journal on Recent Innovation, 3(2), 208-216.

[42] Amin, J., Sharif, M., Gul, N., Kadry, S., & Chakraborty, C. (2022). Quantum machine learning architecture for COVID-19 classification based on synthetic data generation using conditional adversarial neural network. Cognitive Computation, 14(5), 1677-1688.

[43] Misgar, M. M., Mushtaq, F., Khurana, S. S., & Kumar, M. (2023). Recognition of offline handwritten Urdu characters using RNN and LSTM models. Multimedia Tools and Applications, 82(2), 2053-2076.

[44] Singh, V. K., & Kolekar, M. H. (2022). Deep learning empowered COVID-19 diagnosis using chest CT scan images for collaborative edge-cloud computing platform. Multimedia Tools and Applications, 81(1), 3-30.

[45] Bharati, S., Podder, P., Thanh, D. N. H., & Prasath, V. S. (2022). Dementia classification using MR imaging and clinical data with voting based machine learning models. Multimedia Tools and Applications, 81(18), 25971-25992.