IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, March 2023
Impact Factor: 7.301

Form Based Document Understanding using Sequential Model

Khushi Rathi[1], Aniket Patil[2], Nikita Mahadik[3], Sakshi Ahire[4], Abhilasha Shinde[5]
Students, Department of Information Technology Engineering[1,2,3,4]
Asst. Professor, Department of Information Technology[5]
Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract: *DocFormer is a teachable transformer-based multi-modular start-to-finish model for various Visual Report Figuring out projects. Additionally, Doc Former facilitates the model's ability to link text and visual tokens by sharing learned spatial across modalities. For the study of various algorithms and computer vision, it is a crucial area of research. Some people fabricate documents and engage in illegal activities because the current document detection system is ineffective. In the proposed system, fake documents can be found in two ways. As the popularity of deep learning technology has grown, document AI, which includes things like document layout analysis, visual information extraction, document visual question answering, document image classification, and so on, has made significant progress in recent years. We are primarily focusing on the factors that contributed to crimes in previous years when it comes to the factors that contribute to the occurrence of crimes, such as the identity of the offender and other factors.*

Keywords: Crime Analyze, OCR , Feature Extraction, SVM, Accuracy.

I. INTRODUCTION

Form documents often have more complex layouts with structured objects such as tables, columns, and text blocks. Crimes are social nuisances and have a direct impact on society. Governments spend a lot of money to prevent crime through law enforcement. Today, many law enforcement agencies have large amounts of crime-related data that must be processed in order to be transformed into useful information. The proposed system can detect whether a document is genuine and efficient. The image processing combination works very efficiently and the results obtained are accurate. The system is trained on counterfeit documents and how to prevent them using a word processor. Machine learning applications learn from input data and use automated optimization techniques to continuously improve output accuracy. Sequence Modelling is the ability of a computer program to model, interpret, make predictions about or generate any type of sequential data, such as audio, text etc.

It is inappropriate to use a sequential model when:

- Any of your layers has multiple inputs or outputs
- Your model has multiple inputs or outputs
- You need to do layer sharing
- You want non-linear topology (e.g. a residual connection, a multi-branch model)

Deep learning, which is a subset of machine learning, is essentially a neural network with three or more layers. Although they are far from matching the capabilities of the human brain, these neural networks give it the ability to "learn" from a large amount of data and make an effort to imitate its behavior. Although more hidden layers can help improve accuracy, a neural network with just one layer can still make rough predictions. Deep learning makes it possible for numerous artificial intelligence (AI) applications and services to improve automation by carrying out physical and analytical tasks without the need for human intervention. Deep learning technology is behind everyday products and services like digital assistants, voice-activated TV remotes, and credit card fraud detection.

There was definitely need of building such system will help to decrease the work at technician who does data entry eg. in banking sector , at college admission process etc. Along with that this system is going to detect the criminal record as well as the provided document is fake or not. there was no any previous system which was providing different services in a system. Intoducing such system will reduce time at any type of data entry which are related to document or form.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, March 2023
Impact Factor: 7.301

As well as by detection of criminal record of candidate technician can easily debar or accept the form or document of candidate.

II. LITERATURE SURVEY

[1] The paper is based on FormNet: Structural Encoding beyond Sequential Modeling in Form Document Information Extraction. Chen-Yu Lee†, Chun-Liang Li† , Timothy Dozat‡, Vincent Perot‡, Guo- long Su‡ work on this paper. The paper is based on Sequence Die Modeling It showed the state of the art in natural language and document understanding tasks. In practice, however, it is difficult to correctly serialize tokens in documents like forms due to different layout patterns. To mitigate suboptimal serialization of forms, we propose FormNet, a structure-aware sequence model. First, we design Rich Attention, which uses the spatial relationships between tokens in the shape to compute amore accurate attention value. Form Nettherefore explicitly recovers any local syntactic information that may have been lost during serialization. In experiments, FormNet outperforms existing methods with compact model sizes and less pre-training data, establishing new state-of-the-art performance on CORD, FUNSD, and payment benchmarks.

[4]The paper is based on Doc Former: End-to-End Transformer for Document Understanding. Srikar Appalaraju, Bhavan Jasani, Bhargava Urala Kota, Yusheng Xie work on this paper .In this paper , they introduce DocFormer, a multimodal transformer- based architecture for Visual Document Understanding (VDU)tasks. VDU is a difficult problem aimed at understanding documents of various formats (forms, receipts, etc.) and layouts. Additionally, DocFormer is unsupervised and pre- trained using carefully designed tasks to encourage multimodal interactions. DocFormer takes text, vision, and spatial features and combines them with a new multimodal self-awareness layer. DocFormer also shares learned spatial embeddings across modalities, making it easier for models to associate text with visual tokens and vice versa. Doc-Former is evaluated on four different datasets, each with a strong baseline. DocFormer achieves state-of-the-art results on all of them, in some cases even better than four times the model size (number of parameters).

[6]The paper is based on Document Vector Extension for Documents Classification. The author that works on this paper are Shun Guo, Nianmin Yao. The paper is based on Simple linear models, which typically learn word-level representations that are later combined into document representations, have shown impressive performance recently. To improve the performance of document-level classification, it is important to examine the factors that affect document vector quality. This article proposes the concept of containers and further explores the properties of word and document containers through experiments and theoretical demonstrations. Document containers have afixed capacity, and document vectors obtained by simple intersection of too many word embeddings are not fully loaded by the container, resulting in loss of semantic and syntactic information for very large text datasets. It is clear. We also propose an efficient approach to representation of document by using a clustering algorithm to divide the document container into multiple sub containers and establish relationships between the subcontainers.

[3]The paper is based on document ai: benchmarks, models and applications .Lei Cui, Yiheng Xu, TengchaoLv, Furu Wei works on this paper . the paper is based on Document AI, or Document Intelligence, is a relatively new research topic related to technologies for automatically reading, understanding, and analyzing business documents. This is an important research direction in natural language processing and computer vision. In recent years, the spread of deep learning technology has greatly accelerated the development of document AI. B. Document layout analysis, visual information extraction, visual document question answering, document image classification, etc. This white paper provides a brief overview of some representative models, tasks, and benchmark datasets. In addition, we introduce early-stage heuristic rule-based document analysis, statistical machine learning algorithms, and deep his learning approaches, especially pre-training methods. Finally, we will look at the future direction of document AI research.

[7]The paper is based on Fraudulent Detection in Credit Card System Using SVM Decision Tree. Vijayshree B. Nipane, Poonam S. Kalinge, Dipali Vidhate, Kunal War, Bhagyashree P. Deshpande works on this paper With the advancement of the field of e- commerce, fraud has spread all over the world and caused enormous economic losses. In the current scenario, credit card fraud is the leading cause of economic losses. It affects not only professionals, but also individual customers. Decision trees, genetic algorithms, meta-learning strategies, neural networks, and HMMs have been presented as credit card fraud detection methods. Intended fraud detection systems use the artificial intelligence concept

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, March 2023
Impact Factor: 7.301

of support vector machines (SVM) decision trees to solve problems. Therefore, implementing this hybrid approach can further reduce financial losses.

[2] The paper's foundation is Fake Education Document Detection Using Image Processing and Deep Learning. Mrs. G. Chandra Praba, E. Jeevitha, and A. Abitha are the authors of this paper. The paper is based on the fact that official document forgery is becoming increasingly common, posing numerous challenges for public institutions. It is now simple to edit a scanned document and create a new one with different information that is difficult to distinguish from the original and the fake thanks to modern, powerful digital printers and a plethora of software tools. Some people fabricate documents and engage in illegal activities because the current document detection system is ineffective. One of the two methods included in the proposed system for identifying counterfeit documents is the QR-code scanner, which scans the QR code of the document to determine whether it is genuine or counterfeit. Second, there are three stages in the image processing techniques: a phase of testing, a phase of classification, and a phase of training to identify fake documents. The originality of a document is the focus of this proposed project, which focuses on making document forgery detection more reliable and robust. A neural network is used to identify a fake document by combining an error value analysis algorithm with an image processing system.

[5]The paper is based on Hindi Multi-document Word Cloud based Summarization through Unsupervised Learning. Prafulla B. Bafna, Jatinderkumar R. Saini works on this paper .Document management is a critical and important task, supporting a variety of applications from searching for information to clustering search engine results. The multilingual features offered by the website have made Hindi a leading language in the field of digital information technology today. This work focuses on document management and summarization of the Hindi corpus. The goal is to manage documents and summarize Hindi corpus by extracting tokens and grouping documents. This work excels in terms of scalability and supports consistent cluster quality for incremental datasets. Most past and present research has focused on the management of English corpus documents. The Hindi corpus was mainly used by researchers to investigate word stemming, single-document summarization, and classifier design in the Hindi corpus. Word Cloud implementation of unsupervised learning on Hindi corpus for summarizing multiple documents is still unexplored territory.

III. METHODOLOGY

This framework acknowledges datasets of pictures and text as information. We know that we are training datasets and processing data on the system, so we use modules: Preprocessing, feature extraction, and classification all make use of our OCR/SVM algorithm.

The obtained data is first pre-processed with the machine learning technique filter and wrapper to get rid of repeated and irrelevant values. The data have been cleaned because it also makes the data less dimensional. After that, the data are processed further using a splitting process. It is divided into two groups: a trained data set and a test data set. The model is trained using the testing and training datasets, followed by mapping. An integer is used to map the crime type, year, month, location, and time to make classification easier. The free effect between the qualities are poor down at first by using SVM is used for describing the independent features isolated. It is possible to examine when and where a crime occurred because the features of the crime are labeled. Finally, data on the most common crimes are gathered, both spatially and temporally. The performance of the prediction model is determined by examining the document's authenticity and the accuracy rate.

After the dataset has been preprocessed (the pre-processing step cleans the image and removes blur), the system extracts the parameters or features of the dataset in the extraction section. Classification is the next step, where our OCR/SVM algorithm is used.

IV. CONCLUSION

In this paper, we present Doc-Former, a multimodal, continuously trainable, transformer- based model for various tasks of visual document comprehension. A systematic approach to identifying crime is crime analysis and prediction. The proposed system can recognize whether a document is genuine and efficient. Whether the system is trained on forged documents. This project has vast scope. It is possible to build such system for different organization according to their required fields. Also, it is possible to fill information which we get from document directly into the field. We can also provide some features like accepting multiple language will provides flexibility to users as well as technician to extract

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, March 2023
Impact Factor: 7.301

data from document which will be in various languages and building such system for different organization will reduce time of data entry.

REFERENCES

[1]. Chen-Yu Lee, Chun- Liang Li, Timothy Dozat, Vincent Perot, GuolongSu, Nan Hua, Joshua Ainslie, Renshen Wang, Yasuhisa Fujii, Tomas Pfister. FormNet: Structural Encoding beyond Sequential in form document Information Extraction-2022

[2]. Mrs G Chandra Prabha, E. Jeevitha, B. Shwetha .Image processing and deep learning-based document detection for fake education – 2021

[3]. Lei Cui, Yiheng Xu, TengchaoLv, Furu Wei. DOCUMENT AI: BENCHMARKS, MODELS AND APPLICATIONS - 2021

[4]. SrikarAppalaraju , Bhavan Jasani , Bhargava Urala Kota. DocFormer: End-to-End Transformer for Document Understanding -2021

[5]. M.NareshBabu ;S.N.Vyshnavi , S.Keerthi , P. DivyaSree , D. S. V. Naga Prasad .Crime type and occurence prediction using Machine learning algorithm – 2022

[6]. B. Sivanagaleela, S. Rajesh .Fuzzy C-Means Algorithm for Crime Prediction and Analysis in 2019

[7]. Vijayshree B. Nipane, Poonam S. Kalinge, DipaliVidhate, Kunal War, Bhagyashree P. Deshpande.

[8]. Fraudulent Detection in Credit Card System Using SVM & Decision Tree – 2016