Data Glacier

Neural Network for Hand-Written Recognition

Submitted by-

Name – Deekshant Nandeshwar

Email – deekshantnandeshwar@gmail.com

Country – United Kingdom

Internship Batch: LISUM11

Specialization – Master of Science (Data Science)

Content

Abstract

The aim of this assignment is made a report on neural networks for handwriting recognition.

The objective of this assignment is going through all the histories of neural network and various case studies then write a report on that way.

The goal of this report is to made a conclusion by observing all the scenario from history of which neural network is used along with why did they choose those neural network compare to other also what is negative point related to that neural network that made them to create the new neural network which means what neural network comes after that I.e present and how that neural network is better than previous one . Also I have to go through all the case studies from choose the best of them and made a conclusion that how did they overtake the situation that arise and I have to go through the historical survey and write about that how did they preserve the historical document and what method they used for that and what recognition rate are currently attainable.

According to Investopedia neural network is something which made of series algorithms that strive the correlation between data through a process which is as it is copy of the neurons in the brain of human. Like in human brain it learn from our past experience , I want to apply same thing in artificial neural network.

Depend upon the categories of learning so many methods are present which is related to that category which include supervised, semi- supervised and un –supervised. And by performing such kind of method I can achieve state of the art performance when that method is compared with Traditional machine like machine translation , robotic etc. Despite their well-known flaws, most systems depend on the same hidden Markov models that have been used in speech and handwriting recognition for decades. This report discusses an alternative method based on recurrent neural network that was created specifically for sequence labelling tasks with difficult-to-segment data and lengthy, bi - directional, or multi - directional interdependencies.

This report present survey on how neural network transform with time , There are lots of artificial neural networks are present but I am primarily focus on those neural network who can work efficiently on handwritten recognition along with that I am focus on neural network who have highest number of accuracy when it come to pattern recognition from multilayer perceptron used in OCR to CNN used In dynamic learning which bring the new evolution in the field of AI with accuracy 99.87 %. Even the best current recognizers have low recognition rates due to the complexity of overlapping characters, also I used hybrid neural network like RCNN, BLSMT etc. which Work perfectly fine in some situation.

When samples is observed and found that if it is misclassified , it is very hard to recognise even with human eyes and truth is most of them are related to handwritten digits . Futhermore, deep learning explored and evaluated in different domains are also included in this report.

1. Introduction

According to ieee , A neural network (NN) is a computational architecture that consists of a large number of adaptive "neural" processors that are linked in a massively parallel manner. Because of its parallel nature, it can do computations quicker than standard techniques. Because of its adaptive nature, it can react to data changes and learn the characteristics of the input signal. In a NN, there are many nodes. The output of one node in the network is fed to another, and the final decision is made based on the complex relationship between the two nodes.

According to Wikipedia, Handwriting recognition (HWR), also known as Handwritten Text Recognition (HTR), is the capacity of a computer to recognise and translate easily understandable handwritten input from a variety of sources, including paper notes, images, touch screens, and other devices.

In handwriting recognition, On-line and off-line handwriting recognition are the two types of recognition that are commonly used. A time-ordered series of coordinates that represents in on-line recognition. The action of the pen-tip is registered, while only a snapshot of the text is possible when working off-line. Online detection normally produces improved results due to the simplicity with which related features can be extracted. Some other important distinction is between recognising individual characters or words and recognising complete lines of text. Apparently, the other is even more difficult, and the outstanding results achieved for digit and character recognition, for example, haven't been balanced for full blocks. Also, Handwriting acceptance can be divided into two types of situations: those in which the writing style is restricted in any manner (for example, only hand-printed characters are permitted) and those in which it is not restricted. Despite more than four decades of research into handwriting recognition, creating an accurate, general-purpose device for without limit text line recognition remains a challenge.

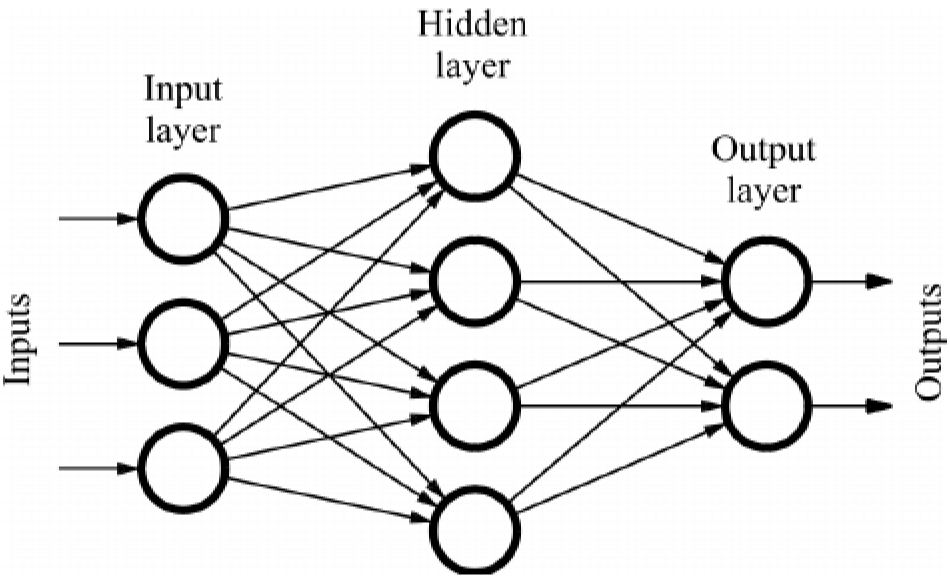
2.1.1 Offline handwritten recognition

In offline handwriting recognition The document is checked after it has been written. The only detail that can be examined is a character's binary output against a backdrop. Although switching to a digital stylus for writing offers more information about pen strokes, friction, and writing speed, offline methods are still needed when the internet is inaccessible. It's particularly important for historical archives, libraries, and large-scale handwritten type digitisation.

History of offline handwriting recognition begin with OCR in 1974 which is developed by Ray Kurzweil. The method was more accurate since the problem domain was reduced. This allowed handwriting forms to be recognised. It lacked efficiency and understanding of unlikely characters, first and foremost.

1) Neural network used for OCR is multilayer perceptron :-

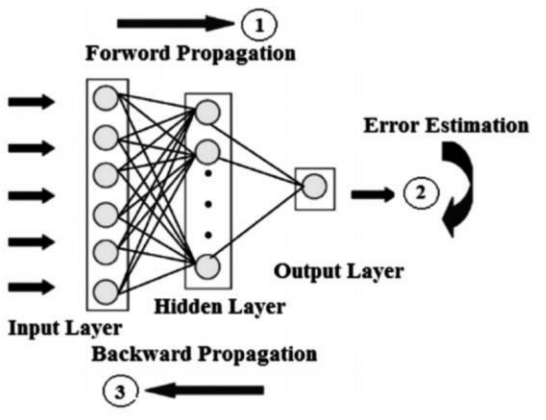
Feed forward neural networks: - feed forward neural network or multi-layer perceptron are the typical deep learning models. Models are called feedforward because information travel through the function is being appraise from a, then the in the middle the computations is used to define f, and then finally there is an output b (information flows always in one direction). when the feed forward neural network extended in such a way that it will feed back the information into itself then it is called as the recurrent neural networks.



Multilayer Perceptron: - a set of instructions designed to perform a specific task which is further by binary classifier for supervised learning. Well, in Its perceptron I will take the input then multiple that input with the weight which results the weighted sum after adding all the weighted multiplied input value. The activation function in perceptron who's job is to create perceptron output by taking weighted sum as a output. Important job of activation function is to ensure that the output should be mapped between the required value.The inclusion of multiple layers, which enable neural networks to model a feature hierarchy, was omitted. As a consequence of the shallow neural network, perceptron was unable to do non-linear sorting, such as the classic logic XOR function.

Multilayer perceptron took place when backpropagation and hidden layers are introduce in perceptron . whereas , Backpropagation is a technique for constantly changing weights in order to reduce the difference between the actual and desired output. Neural networks can learn more complex features thanks to hidden layers, which are nodes of neurons layered between inputs and outputs (such as XOR logic).

MLPs and other feedforward networks are similar to ping-pong they primarily involved in two movements, a constant back and forth (forward and backward passes)



Advantage

fast (highly parallel)

simple to understand

capable of partitioning the input feature space into arbitrary partitions

Disadvantage:

The learning time of multi-layer perceptron networks with backpropagation scales exponentially for complex Boolean functions as small training sets are used.

Application

MLPs are useful in science because they can solve problems in a stochastic manner, allowing for approximate solutions to extremely complex problems such as fitness approximation.

MLPs are a basic feature that can be used to construct statistical models using regression analysis, as well as in a variety of other fields including image recognition.

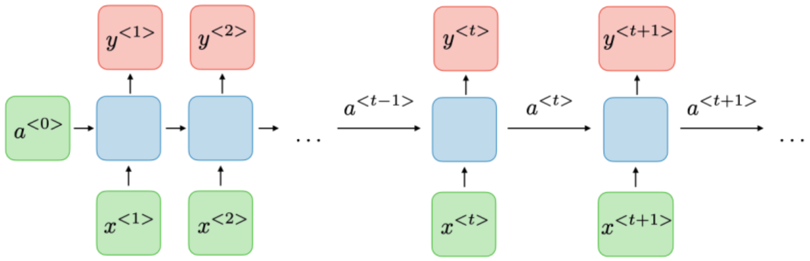
- Recognition rate of multi level perceptron on test set is 96.7%.

2.1.2. Online handwritten recognition

The automated translation of text as it is written on a special digitizer or PDA, where a sensor picks up pen-tip motions as well as pen-up/pen-down switching, is known as online handwriting recognition. Digital ink is a type of data that can be thought of as a digital representation of handwriting. The signal is encoded into letter codes that can be used in computer and text-processing programmes.

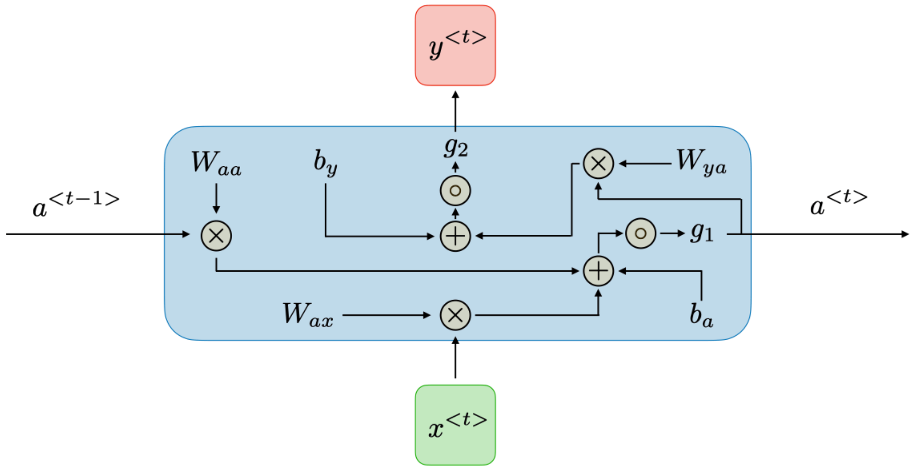
According to wikipedia The aim of preprocessing is to exclude all redundant information from the input data that could interfere with identification. This has to do with speed and accuracy. Binarization, normalisation, sampling, smoothing, and denoising are popular preprocessing techniques. The extraction of features is the next step. Higher-dimensional data is derived from the two- or higher-dimensional vector fields received from the preprocessing algorithms. The classification process is the final major move. Various models are used in this process to map the extracted features to different classes, allowing the characters or terms the features represent to be identified.

a) Recurrent Neural Networks: - RNNs, or recurrent neural networks which introduced in 1986, are a kind of neural network which has hidden states which allow former outputs being used as inputs. They usually go like this:



the activation a^{< t >}a and and the output y^{< t >}y are represented in the following way for each timestep tt:

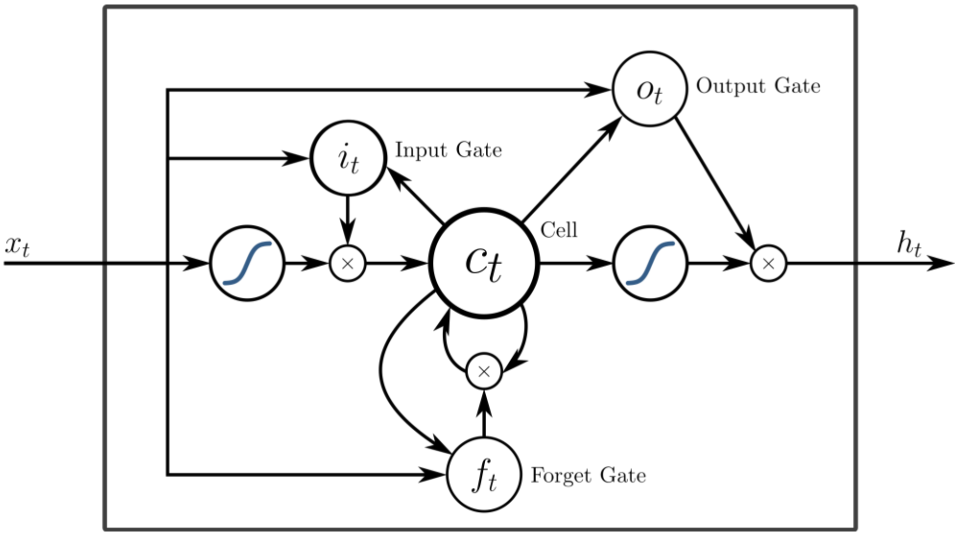
Wax,Waa,Wya,ba,by are **exchanged** temporally coefficients and activation functions g1, g2



Using their internal state, RNNs will process variable length sequences of inputs (memory). As a consequence, tasks like unsegmented, connected handwriting recognition are possible.

To solve problems with long-term temporal dependencies, standard RNNs can be difficult to train. That’s is reason in 1997 LSTM is introduced in 1997.

Long/short term memory :- A memory cell is introduced in this form,and is a special cell that can handle data with time gaps (or lags). RNNs will process texts by “remembering” ten prior sentences, while LSTM networks can process video frames by “remembering” anything that occurred several frames before. Writing and speech recognition are two other applications of LSTM networks.



The input gate determines how much information from the previous sample will be stored in memory; the output gate controls the amount of data passed to the next layer; and the forget gates control the memory tearing rate.

Application

Use in language translation

Used in speech and hand written recognition

Disadvantage :

The issue is that the data must also be moved from cell to cell in order to be evaluated. Furthermore, with the inclusion of new features to the frame, the cell has become very complex.

Different random weight initializations affect LSTMs, making them behave similarly to a feed-forward neural network. Instead, they choose minimal weight initializations.

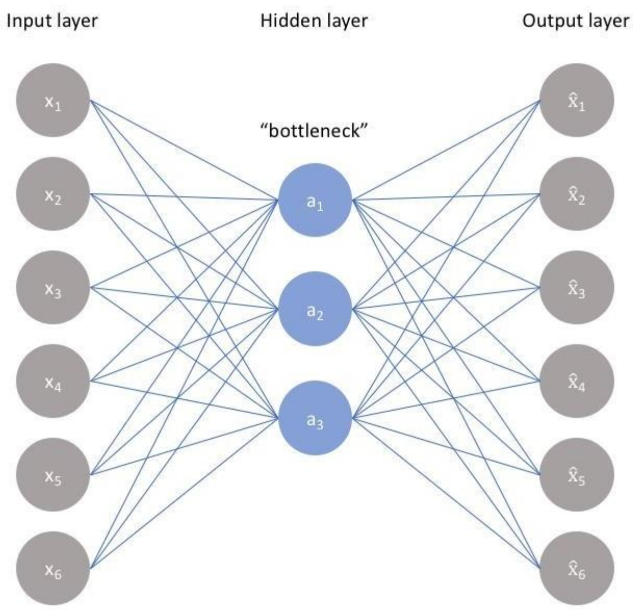
When it comes to training both rnn and lstm, connectionist temporal classification (CTC) plays important roles.

connectionist temporal classification (CTC)

According to wikipedia Connectionist temporal classification (CTC) is a form of neural network output and related scoring feature that can be used to train recurrent neural networks (RNNs) including LSTM networks to solve sequence problems with variable timing. It can be used for online handwriting recognition, for instance.

b) Autoencoders :-

For sorting, clustering, and attribute compression, autoencoders are used. When training FF neural networks for classification, you usually feed them X examples from Y categories and expect one of the Y output cells to be triggered. This is referred to as "supervised learning." In the other side, without guidance, they may be prepared. Their arrangement forces AEs to generalise data and look for common patterns when the number of hidden cells is less than the number of input cells (and the number of output cells equals the number of input cells), and when the AE is trained in such a way that the output is as similar to the input as possible.



Application :

I can use it for Minimization of Dimensionality

When it comes to image compression and Binarization an image autoencoder work very well.

Advantage

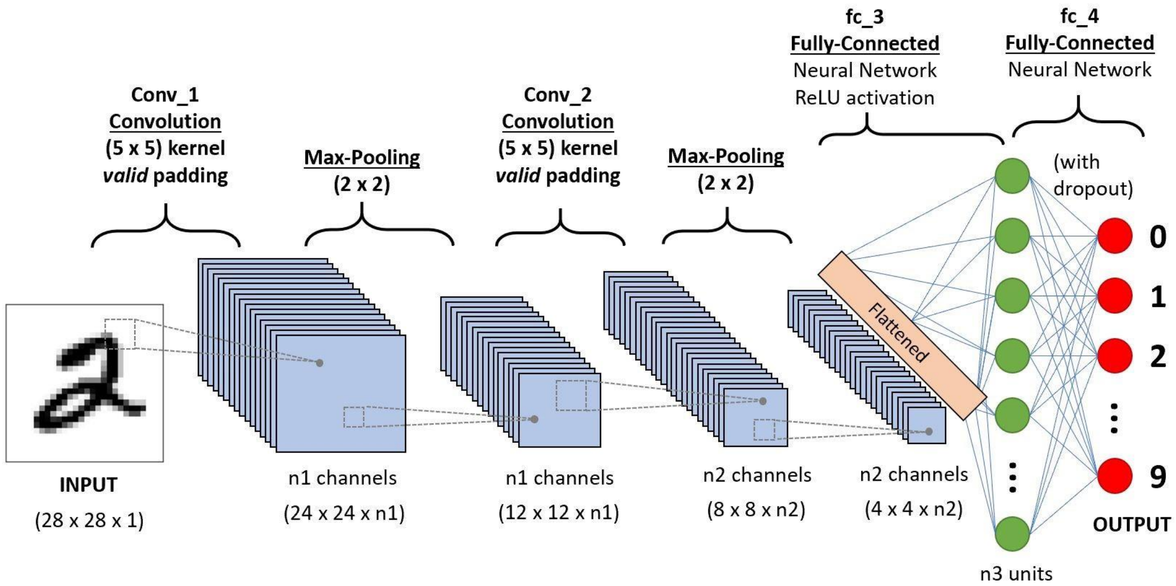
Autoencoders are useful for extracting features.

Disadvantage

If your training data does not match your testing data, you end up misrepresenting rather than clarifying details.

c) Convolutional Neural Network:

Convolutional Neural Network comes into the pictures in 2010 , CNN is a Deep Learning algorithm that can take an input image and attribute significance (learnable weights and biases) to different aspects/objects in the image, as well as discriminate between them. When compared to other classification algorithms, a CNN needs considerably less pre-processing. CNN can master these filters/characteristics with enough practise, even though they are hand-engineered in simplistic ways.CNN is similar to the pattern in the neuron in human brain .



Application :

Text analysis can be done with convolutional neural networks. This is not only beneficial for handwriting research, but it also has a significant impact on recognizers. In order for a computer to search an individual's writing and equate it to a large archive, it must run nearly a million commands per minute. According to reports, the error rate at the character level has been reduced to a minimum of 0.4 percent thanks to the use of CNNs.

Natural history libraries, for example, use CNNs for more complex reasons. These collections play an important role in recording global historical events such as ecology, evolution, habitat destruction, biological invasion, and climate change, among others.

Limitation :

The location and direction of the target are not encoded in CNN's forecasts. They totally forget all internal details about the object's pose and direction, and they send all of the information to the same neurons, which may or may not be able to handle it.

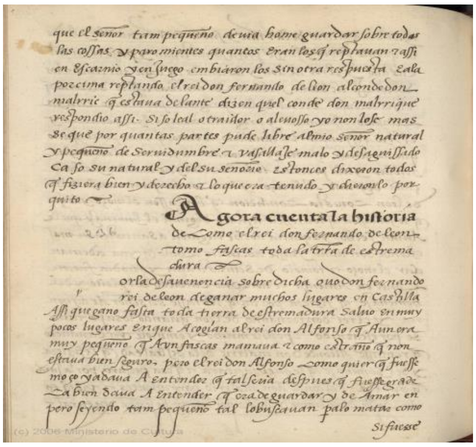
Max pooling discards important data and fails to encode relative spatial relationships between functions. As a result, CNN are not necessarily invariant to massive input data transformations.

CNN's identification accuracy for an MNIST dataset was 99.87 percent.

2. Historical survey

The digitization of ancient handwritten paper images is needed for the preservation of cultural heritage. Furthermore, transcription of text images acquired by digitization is needed to provide productive information access to the content of these documents.Handwritten Text Recognition (HTR), It's become a hot topic in image and computational language processing to obtain summaries from text files. In the other hand, modern HTR systems are far from optimal. One challenge is dealing with image distortion and handwriting inconsistency in ancient historical texts, as well as a vast number of Out-Of-Vocabulary terms. One solution to this problem is to use external lexical methods. The aim is to combine a good optical recognition system which can interact with image distortion and uncertainty with a sub-lexical unit-based language model that can model OOV terms Experiments with a recognizer built on Hidden Markov Models, such as digitising an ancient Spanish manuscript . This method is then used to classify deep network classifiers such as (BLSTMs) and (CRNNs)

Handwritten text recogniser (HTR) tools, on the other hand, face a number of difficulties in terms of image and language modelling. Due to age, fabrication, and ink spill, historical record photographs can contain flaws. They can also contain decorations such as calligraphic initial letters and broad character strokes.example are :-



There are so many methods for developing optical models for handwriting recognition have been suggested.. Long Short-Term Memory (LSTMs) and other Hidden Markov Models (HMMs) and Recurrent Neural Networks (RNNs) and their variants: Examples of such approaches include Bi-Directional LSTMs and Multi-Dimensional LSTMs . As inputs, HMMs and RNNs accept sets of components traditional made by hands . Convolutional layers are the first layer of Deep learning methods when it start and also it focus on extracting features based on learning not component traditionally made by hands.

Optical models are often consistent with lexical models and Language Models at the word level in HTR systems to guide the identification of individual words and possible word sequences. Accodring to the Rodirgo corpus text, for example, contains a significant variety of OOV words and many unusual phrases and words in their ancient forms. Additionally, this corpus contains OOV characters that are not used in the training set (such as: \, ´p, ¯g, } and w). OOV terms include acronyms, words hyphenated differently in the training and test sets, and terms which occur in various forms in the training and test sets (for example, portugal and portu gl). The sub-word unit language model is trained using the transcription of the testing partition's text row after a brief pre-processing phase. This preprocessing entails inserting a new symbol for word slicing, as well as breaking down the words into sub-word sequences. The nuances of the term division are preserved in this manner. The sub-word unit language model can then be trained using these preprocessed transcriptions.

Typically, massive n-gram language systems with sub-word units are trained . The lexicon, on the other hand, has been condensed to accommodate the list of sub-word groups.

To arrive at the final hypothesis, the best hypothesis is processed in the encoding process.

HTK was used to train Hidden Markov Models (HMM) for optical modelling. The qualified models are four-state left-to-right character models. A mixture distribution of 64 Gaussians describes the observation probabilities in each state. It uses a stack of 13 convolutional layers (3 3 filters, 1 1 stride) along with Three bi-directional LSTM layers of 256 units each followin a Convolutional Recurrent Neural Network.

Deep optical models allow for a statistically meaningful relative improvement of 59.2 percent in WER over the HMM system (43.9 percent 0.5) and an 81.1 percent relative improvement in CER over the HMM system.

3. State of the art

UNIPEN (online handwritten recognition database) testing ground for handwritten character recognition in isolated environments. Cluster generative computational dynamic time warping, a hybrid technique that combines dynamic time warping with HMMs and embeds clustering and statistical series of modelling in a single feature space, has been found to work well on UNIPEN, as has a svm with a novel Gaussian dynamic time warping kernel. On UNIPEN, standard error rates for digit recognition range of about 3% and 10% for lower case character recognition. Similar approaches can be used to classify isolated words, and this has shown promising results with small vocabularies (4.5% error rate for 32 words). For general vocabulary recognition operations, specific characters are identified and afterwards mapped into complete words using a dictionary. I might naively achieve this by segmenting words into characters and categorising each line separately. For cursive or not limited text, segmentation is difficult until the phrases have already been recognised. However, approaches have been proposed in which segmentation is performed first, followed by recognition. Start with basic strokes instead of characters to segment the letter. Velocity minima, y-coordinate minima, and maximum curvature points are some of the ways the stroke boundaries can be defined. Sayre's (The circular dependency between segmentation and recognition is referred to as Sayre's paradox.) dilemma may be solved by segmenting and recognising at the same time. One of the reasons hidden Markov models (HMMs) are so common for unregulated handwriting is because of this. HMMs are considered to have a variety of defects. One of them is that they believe that the likelihood of any observation is purely determined by its current state, making it impossible to model contextual results. HMMs often have the advantage of being generative, while discriminative systems are better at labelling and classification functions.

Recurrent neural networks (RNNs) seem to be a good alternative to HMMs because they don't have these disadvantages. Nevertheless, the use of RNNs for handwriting recognition has been limited to independent character recognition up until now. Mixing neural networks and HMMs in a so-called hybrid approach has proved to be a more powerful neural network application for handwriting recognition. Although hybrid models make it easy to apply context to HMMs, they also have many of the same shortcomings as HMMs and don't fully use RNNs' sequence modelling capabilities.

4. Case studies

5.1. On Urdu handwritten recognition

One of the most discussed classification problems is handwriting recognition. Handwriting recognition programme converts handwriting script into machine-readable code and can be used both offline and online. Because of the writer's particular preferences in drawing character forms and combining different characters, handwritten text is found more difficult to recognise than typed text.

Several experiments have looked at the identification of Arabic handwriting from the viewpoint of cursive text recognition schemes (such as “Arabic, Urdu, and so on”).

Deep learning-based approaches have been widely used to recognise offline and online Arabic handwriting, with recent contributions reporting high recognition rates. Such similar cursive languages, such as Persian, Pashto,urdu and others, did not receive the same amount of attention as Arabic in terms of recognition schemes.

This case study concentrates on the Urdu script, which is one of many Arabic-derived cursive scripts. More than 100 million people speak Urdu around the world, with the majority hailing from countries from Asia and other continent . The Urdu alphabet is a superset of language spoken by middle east countries, and the language is affected by “persian and arabic ”.

The methodology relies on implicit character segmentation and uses text lines for recognition. Each line of texts is interpreted as a collection of strokes. A convolutional neural network derives machine-learned features from these text lines, and a bi-directional long short-term memory (LSTM) network categorises the feature sequences.

Holistic and observational mechanisms are two common classifications for Urdu recognition approaches. Although holistic techniques in Roman scripts allude to appreciation at the word level, In urdu and arabic , incomplete words or ligatures are referred to. Analytical methods, on either side, lead to understanding at the character stage. Ligatures are used as identification units in holistic processes. When one or more characters have been joined together then it became ligature. Whereas ligature can also classified into Primary and Secondary elements, with primary elements representing the ligature's main section and secondary elements representing the dots and diacritics. In Urdu, there are over a hundred distinct ligatures. Typically, the problem is solved by considering small ligatures or separating the ligatures into two :- primary and secondary and then bring together again in the last stage.

Instead of segmenting ligatures into characters, which is a daunting task in and of itself, the proposed analytical recognition strategies for In Urdu text, segmentation in advance of characters are assumed, or recognition and segmentation of character at same time is used.

Though , recently recognition of text printed is improved, there is still a paucity of research on handwritten text recognition. This thesis used a support vector machine classifier to characterise a mixture of gradient and structural characteristics. The CENPARMI Urdu word index was used to assess the device, and it earned high scores for character recognition.

Handwritten images are divided into text blocks, which are then traversed with fixed-size sliding windows to retrieve pixel values as functions. Such image pixels are fed into a RNN, which help the RNN so that it will learn the ground truth transcription and character type. Just for experiment on UCOM dataset , I took a small part from from it and it produced a character recognition score of 94 percent. Later, the study was expanded and tested on a greater number of samples, yielding the recognition score of 92 percent .

Tactic segmentation-based methods are more effective, according to a detailed review of Urdu text recognition systems. Though further training sets are needed, those techniques avoid the time-consuming overt segmentation stage and perform on Holistic methods which have greater number of distinct groups.

Handwriting consists of a series of strokes that must be converted into proper translation. As a result, recurrent nets are a popular form of handwriting recognition. Recurrent network efficacy has been demonstrated using printed Urdu text, in which pixel values or statistical features are fed into the network to understand segmentation points and shape of the character through window sliding across lines of text. To take it a step forward, raw pixel values can be replaced with CNN-derived computer-trained functions. For classification CNNs are passes the result which it obtained, to the recurrent layer, and this hybrid mode of CNN and LSTM has been shown to be a good method for recognition tasks.

In preprocessing step, source handwriting image is binarized, the size of the text line is normalised, and image obtained is inserted into the convolutional layer. CNN generate large number of feature maps, which are transformed into feature sequences and delivered to the recurrent network through sliding windows.

The CTC layer, After the LSTM layers, the feature sequences are aligned with the ground truth transcription during planning, and the LSTM layer output is decoded during assessment to achieve the desired transcription.

Finally, on a limited range of 50 lines of text in training and 20 in the exam, a recognition rate of 94 percent was registered. On 1840 test lines, LSTMs recognition on raw pixels yielded a 92 percent recognition score. On a dataset of 6000 unique text lines yielded a character recognition score of over 83 percent on average.

Future expectation :-

The aim is to see how well implied segmentation based recognition compares to ligature level recognition. To limit the number of character groups, independent identification of main body ligatures and dots may also be investigated.

5.2. On Devanagari handwritten recognition

The implementation of a methodology for recognising continuous handwritten texts in the same language can be aided by a strong method for recognising individual characters. A handwriting recognition system's overall pipeline includes preprocessing, attribute extraction, identification, and postprocessing. Modules for segmentation and analysis are the most essential components of a competitive recognition scheme. When constructing a robust recognizer, handcrafted components were historically considered in conjunction with one or more state-of-the-art classifiers such as the hidden Markov model (HMM). Since, scientist showed the usefulness of convolutional neural network structure in recognising handwritten English number, also it is observe that there is more complex character recognition problems are waiting for us. So scientist from all over the world switch and started focusing on more than one or combination CNNs. Recognition mechanisms based on convolutional neural networks (CNNs) have the advantage of not needing manual collection of high-level functions. From the original pixel values of image samples, a network like this can extract a valuable range of features. Also apart from this there some more neural network that is used for handwritten recognition and very well on that some of them are :- Recurrent neural networks (RNN) along Long Short Term Memory (LSTM) or Bidirectional Long Short Term Memory (BLSTM) .

The effect of network depth on recognition performance is discussed in this case study, which discusses a significant number of neural architectures were recently studied for offline Devanagari character recognition in a recent report.. The highest level of identifying accuracy was 96.9%, which surpasses the most recent accuracy of character samples of Devanagari (handwritten).

Three main network architectures – a) a combined CNN and BLSTM, (b) BLSTM, and (c) CNN network – were added with varying depths. In the second case, they used a separate node to extract high-level features, and they computed histogram of oriented gradients features to feed into the BLSTM network. There is no other external function extraction module used in the remaining two scenarios.

The HOG attribute descriptor has previously been used for a variety of uses, including scene text recognition. It's calculated by splitting the image into smaller sections, all of them contains a 1-D HOG directions for all pixels .In order to normalise all the smaller areas which belong to the larger area the value of local histogram are distributed over bigger area . And value which obtained after normalised in the larger region is called HOG descriptor A cell or a block refers to the image's smaller and larger regions, respectively. In fact, the block raster scans a portion of the sample image with some overlap. Concatenating the HOG descriptor values results in a high-level attribute vector. A pair of LSTMs known as bidirectional long short term memory (BLSTM) can be used to recall sequences both forward and backward. BLSTM networks have been active in acoustic modelling, handwriting recognition, and other applications. In a BLSTM network, A high-performance recognition method is created by using HOG characteristics derived from character images as a series of attribute vectors.

CNN recognition

A CNN architecture consists of multiple convolution layers accompanied by one or more hidden layers in a completely integrated feedforward network. Within two convolution layers, a subsampling or pooling layer may or may not exist.

A CNN that extracts features from a raw pixel image is called a convolutional neural network (CNN). The amount of attachment weights and the risk of over-fitting have been minimised thanks to a CNN architecture. A CNN's convolutional layer consists of trainable filters. The output required for forward pass of training is nothing but the addition of process of calculating of the input image and multiplication of filter. After that, filter slides and returns the same answer.

A single function value is extracted from each window by the pooling layer. The sum of the entries in each window is often extracted using maximum and average pooling, if the entries is average value. This window slides over the results of the previous layer's convolution method based on the defined parameter. The function vector's dimension is reduced by the pooling position, which excludes the most powerful low-level functions.

According to research paper , a) CNN-BLSTM hybrid architectures (5 depths), b) CNN architectures (5 depths), c) BLSTM network (HOG input features) and d) Multi-layer perceptron (HOG input features) are all examples of multilayer perceptron having HOG input features. Raw images of three different normalised sizes were fed as a input into Convolutional neural network and combination of CNN and BLSTM architectures. And To improve the network generalisation , the Real-time data augmentation is used. A training sample is subject to random deformations such as rotation, spinning, and blurring when it is added to the network in RTDA.

This is the first study of its kind on handwriting recognition using deep architectures of different sizes (a total of 25).

Furthermore, A convolution neural network consist of seven layers and two completely linked layers achieves a high recognition accuracy of 96.09 percent, outperforming Devanagari characters' (offline) latest state of the art accuracy.

A mixed network of five convolutional layers, two bi-directional(LSTM) layers, and two totally linked layers, on other hand, performs poorly.

5. Conclusion

In this report I have provided the knowledge of neural network used and its application over past years. I have reviewed different state of art learning model for different categories. I Have explained the case studies and method it use for recognising the character along the recognition accuracy . I have explain historical survey , how and with what method It be can be used to preserved ancient document and how neural network play important role in that .

6. Reference

Computational intelligence paradigm in advance pattern classification by marek R ogiela and lakhmi C. jain

https://towardsdatascience.com/https-medium-com-rachelwiles-have-we-solved-the-problem-of-handwriting-recognition-712e279f373b

https://www.sciencedirect.com/science/article/abs/pii/S0893608005801443#:~:text=An%20optical%20character%20recognition%20(OCR,of%20the%20input%20feature%20space.

https://www.sciencedirect.com/science/article/abs/pii/016781919090079O

https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks

https://en.wikipedia.org/wiki/Connectionist_temporal_classification

https://stats.stackexchange.com/questions/222584/difference-between-feedback-rnn-and-lstm-gru

https://www.geeksforgeeks.org/understanding-of-lstm-networks/

https://www.researchgate.net/publication/311672974_Offline_Handwriting_Recognition_Using_LSTM_Recurrent_Neural_Networks

https://arxiv.org/ftp/arxiv/papers/1706/1706.06720.pdf

https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53 https://sakhawathsumit.github.io/sumit.log/2018/07/21/drawbacks-of-convolutional-neural-networks.html

https://hal.inria.fr/inria-00104481/

https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464

https://en.wikipedia.org/wiki/Neural_network

https://www.mdpi.com/2313-433X/4/1/15

https://ieeexplore.ieee.org/document/8876916

https://www.researchgate.net/publication/325994503_Does_Deeper_Network_Lead_to_Better_Accuracy_A_Case_Study_on_Handwritten_Devanagari_Characters