The Paper OCR – Technology to Authenticate Genuine Paper

Shrikant Rajendrakumar Deshmane
Student
Department of Master of Computer Application
ASM IMCOST, Thane, Mumbai, India

Abstract— Presently in the modern world, the innovation is going enormous increasing day by day. As much as the innovation builds, the requirement for security is substantially more. In India the data innovation is the best thing to execute security of such significant things. This paper will talk about the usage of OCR method to utilize authenticate watermark and logo of the paper. As the documents are called "legal document" There is some identification or mark technique used to make document legal. And this can be the secured reason to identify the original or genuine document. The OCR technique is used to recognize text printed on document/image where as in India some of us follows watermark to protect our document. The Paper OCR technique is going to develop to authenticate genuine paper or document in order to prevent frauds.

Keywords: OCR, Watermark, Logo, Legal Documents, Stamp Paper, Authentication, Security

I. INTRODUCTION

In the globalized and changed business condition of the most recent couple of years, we face a radically expanding volume of fakes, particularly in the money related segments in India. The Indian budgetary administrations segment has seen exponential development in the most recent decade—a development that has not been without its entanglements, as occurrences of misrepresentation have additionally been on the ascent. Extortion results in huge misfortunes to the general population exchequer, along these lines unfavorably influencing administration conveyance.

Budgetary extortion is enormous business, adding to an expected 20 billion in direct misfortunes every year. Industry specialists speculate that this figure is in reality a lot higher, as firms can't precisely recognize and gauge misfortunes because of misrepresentation. The most exceedingly terrible impact of budgetary fakes is on Financial and the archives inflows into India.

II. EXISTING TECHNIQUE

In the existing scenario we have OCR technique which is use to identify or recognize the character/text from the document and the image.
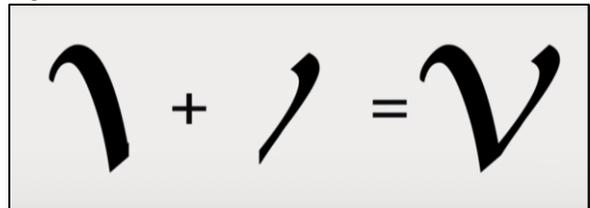
This techniques can be the most powerful to prevent the fraudulent activity against the false documents.
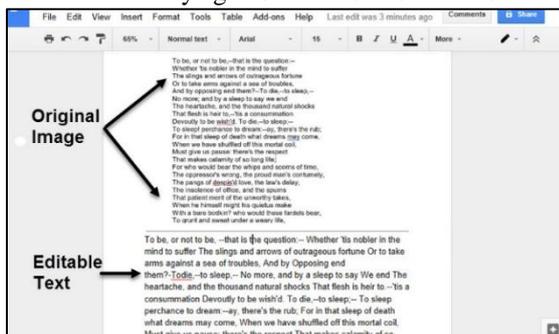
Basically OCR can follow 2 approaches,

− First approach is Future Recognition to identify character by lines or stokes. In this techniques we can identify the character by the strokes and lines of the charterer.

For eg.

− Second approach is Pattern Recognition. In this approach the image is converted into binary digit form to identify the character shape.

For eg:

Basically this approach uses the distance formulae to find the exact co-ordinates as per the single character. Form the binary conversion we get the exact editable character.

III. PROBLEM STATEMENT

There are so many fraudulent cases are registered, Benefits are obtained on the basis of false documents in India.

The motivation behind extortion might be financial addition or different advantages, for instance by getting a visa, travel archive, or driver's permit, or home loan misrepresentation, where the culprit may endeavor to meet all requirements for a home loan by method for false explanations.

To solve this problem we can have a new technique which can authenticate the document or paper with the help of its logo, watermark, stamps, and barcodes, web links.

IV. PROPOSED METHODOLOGY

To overcome the current problem we can have a new techniques which is Paper OCR, which is purely based on the OCR and the OMR.

In this technique we have one example to read a logo from the Paper OCR, the image can be capture by the android phone, it will scan the image and produce its authenticity that what is text written on it.

It follows the some steps include:

− It convert the image to its gray scale,

- Then it is converted to binary form to read the pattern over it.
- Sometimes pattern includes the watermark, which is fade. But this technique can be overcome to this issue.
- Watermark can be read with the help of OMR and OCR which is mixed in algorithm, it will find the exact pattern to identify the logo, watermark.
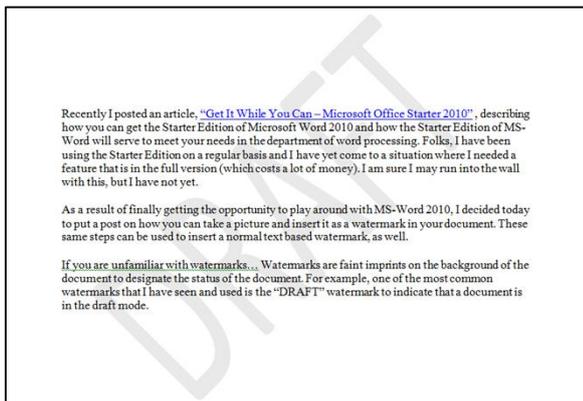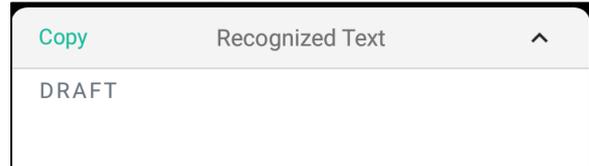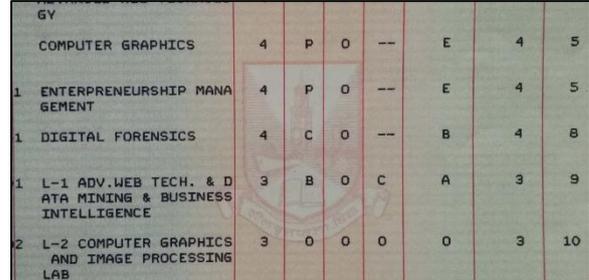
For eg:
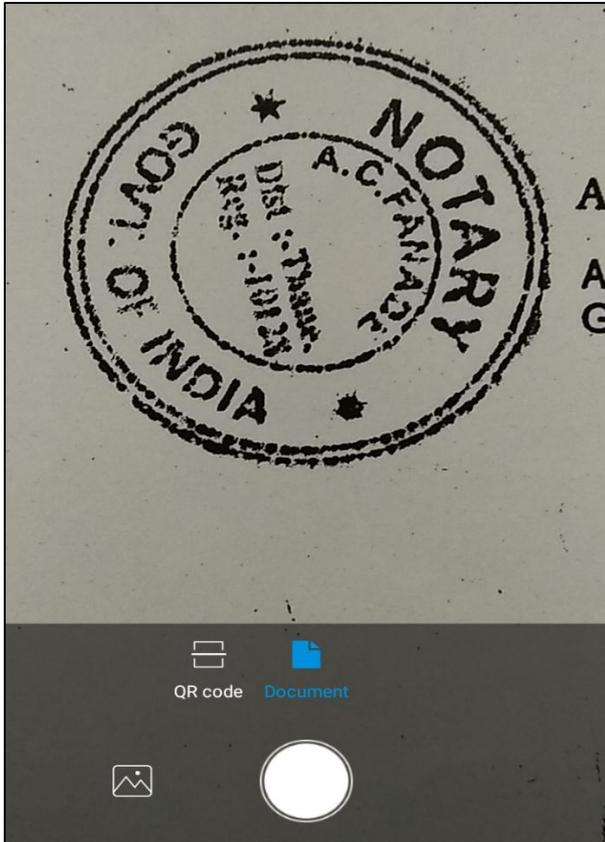
Which will result into:
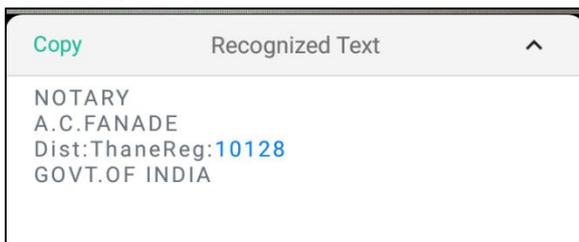
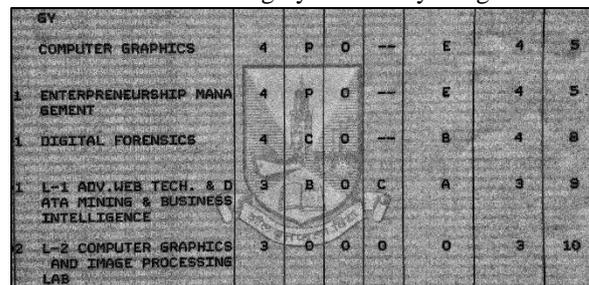Another example is to detect the watermark inside the document.
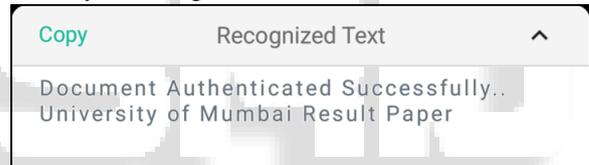

And the result will display:

Now the another example to get document authentication,

First it will converted into gray and binary image.

And finally we will get result into this,

V. RESULT & PERFORMANCE ANALYSIS

According to the philosophy and action, this methods can explain the different issues of deceitful records in India.

The main problem statement indicates the false degrees are currently a billion-pound industry, and cause genuine issues to managers and enrollment specialists, just as possibly risky outcomes to their clients.

For that we can have solution that a portal which helps to identify the authenticity of document as like KYC, the company person will take all the details of the document and authenticate it by using authenticate website.

By such a way the work age should be possible inside the piece of territory.

A. Advantage:

- highly accurate text,
- fast and reliable,
- easy to convert and store data
- user friendly

B. Disadvantage:

- expensive,
- Images takes lots of space
- image quality will be reduce if the paper is too old

VI. CONCLUSION

As per the methodology this technique will help to lead the document as "Legal Document", and can reduces the fraud from the various sectors.

REFERENCES

[1] https://www.slideshare.net/karanpanjwani752/optical-character-recognition-ocr
[2] https://www.idcheck.io/
[3] https://www.futarium.com/blog/2018/12/07/fake-diploma-uk/
[4] https://www.google.com/search?q=how+ocr+works&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiOr56r6b7iAhWCF3IKHeAwCPkQ_AUIDygC&biw=1366&bih=608#imgdii=G1DvVV0kfMnU6M:&imgrc=I5o2bAIz2ebPIM: