veriff

ML Research Engineer

Paper Reading Club

Candidate: Víctor Segura Tirado

Paper: A Multiplexed Network for End-to-End, Multilingual OCR

Index

1. OCR introduction
2. Real-world problem solving
3. Dealing with different languages
4. Potential value

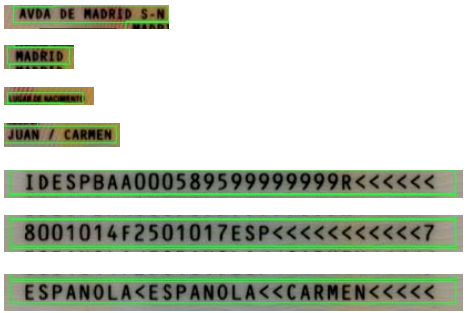
Index

1. **OCR introduction**
2. Real-world problem solving
3. Dealing with different languages
4. Potential value

1 OCR introduction

1. Text detection: detect and label text

back side of spanish identity document



bounding boxes

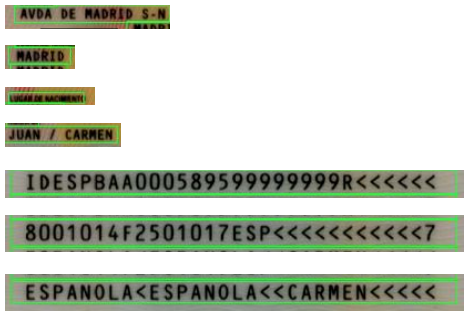
2. Text recognition: convert image into machine-encoded text

{
"street address": "avda de madrid s-n madrid madrid",
"authority": "28391a6dk",
"city address": "madrid madrid",
"parents": "juan carmen",
"mrz": "despbaa000589599999999R<<<<<<
8001014f25010117esp<<<<<<<<<<7
espanola<espanola<<carmen<<<<<<<"
}

1 OCR introduction

1. Text detection: detect and label text

back side of spanish identity document



bounding boxes

2. Text recognition: convert image into machine-encoded text

```
{
    "street address": "avda de madrid s-n madrid madrid",
     "authority": "28391a6dk",
     "city address": "madrid madrid",
     "parents": "juan carmen",
     "mrz": "despbaa000589599999999R<<<<<<
                8001014f25010117esp<<<<<<<<<<7
                espanola<espanola<<carmen<<<<<<<"
}
```

Aplications

identity document verification systems

scanning documents for digital content management

license plate recognition for parking management

Index

1. OCR introduction
2. **Real-world problem solving**
3. Dealing with different languages
4. Potential value

2 Real-world problem solving

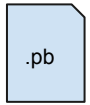
✗ Training a system for each language is **computationally expensive**

N languages = N models

2 Real-world problem solving

✗ Training a system for each language is **computationally expensive**

N languages = N models

✓ **Multilingual approach:** global optimization and high maintainability in real applications → 

N languages = 1 model

✓ **Text Spotting:** text detection and text recognition can be learned jointly with benefits

✓ **Language Predictor Network:** determine a language for text recognition module

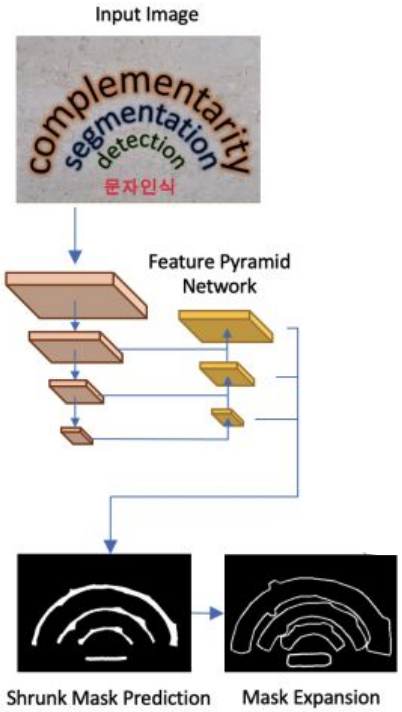
End-to-End framework

Index

1. OCR introduction
2. Real-world problem solving
3. **Dealing with different languages**
4. Potential value

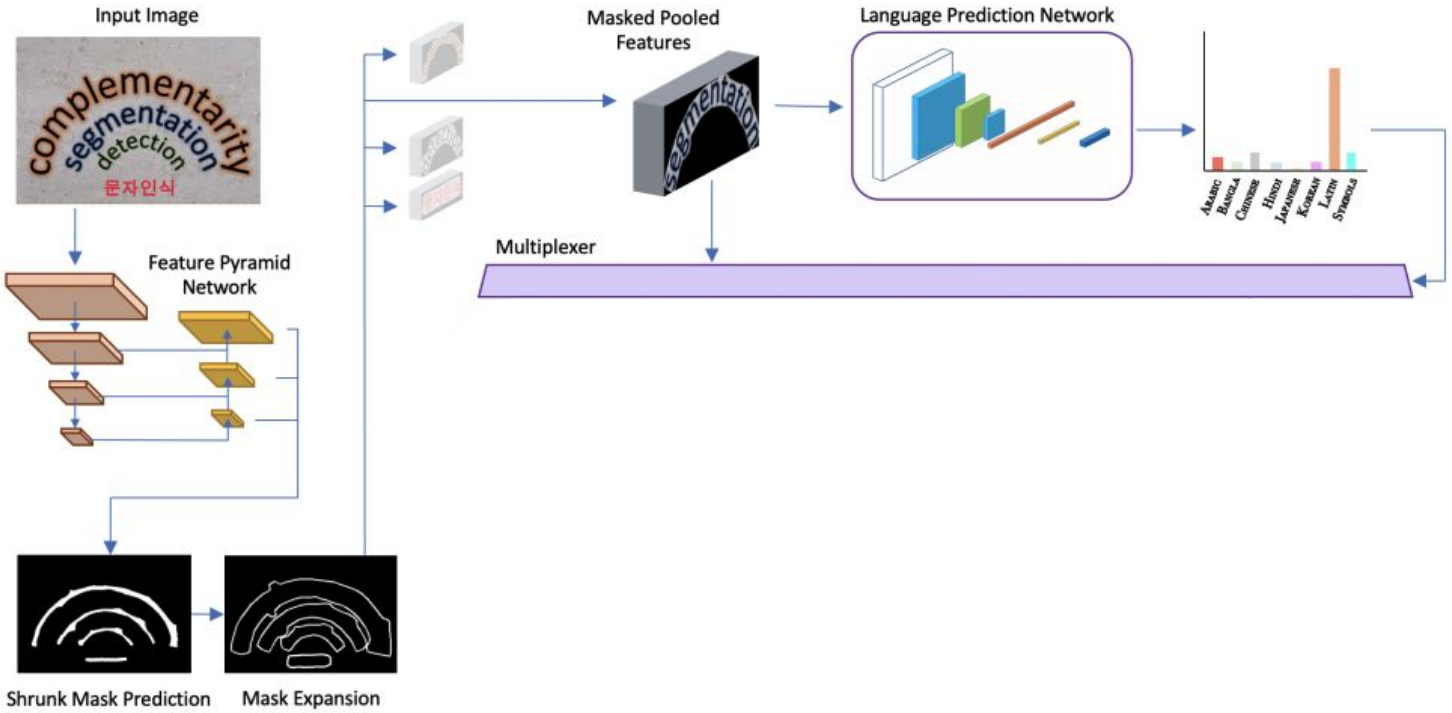
3 Dealing with different languages

M3 TextSpotter: End-to-End OCR capable of recognizing text from multiple languages co-occurring in the same scene



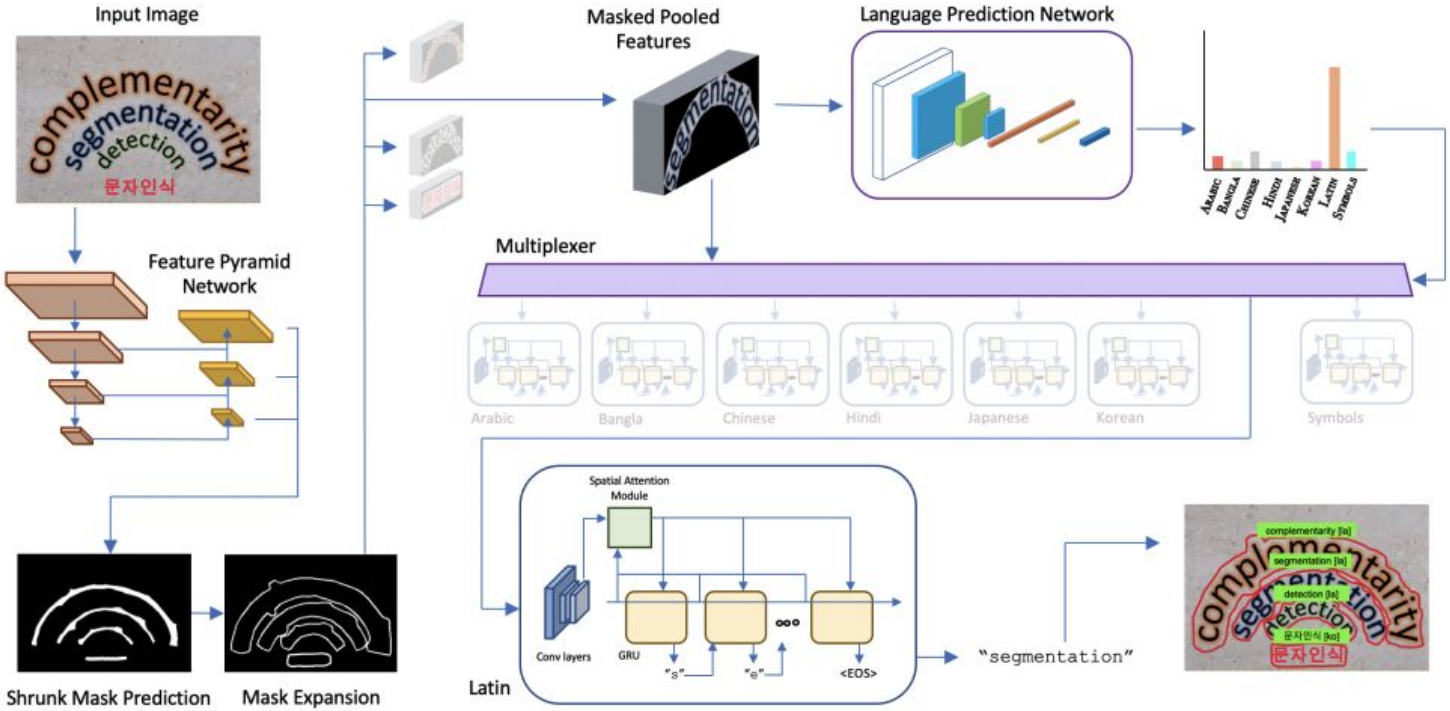
3 Dealing with different languages

M3 TextSpotter: End-to-End OCR capable of recognizing text from multiple languages co-occurring in the same scene



3 Dealing with different languages

M3 TextSpotter: End-to-End OCR capable of recognizing text from multiple languages co-occurring in the same scene

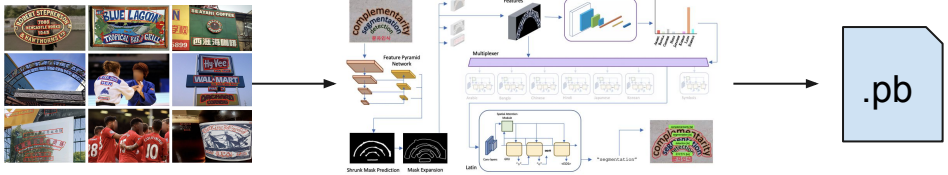


Index

1. OCR introduction
2. Real-world problem solving
3. Dealing with different languages
4. **Potential value**

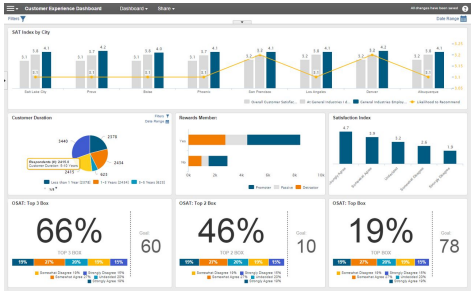
4 Potential value

End-to-End framework that trains in an end-to-end manner



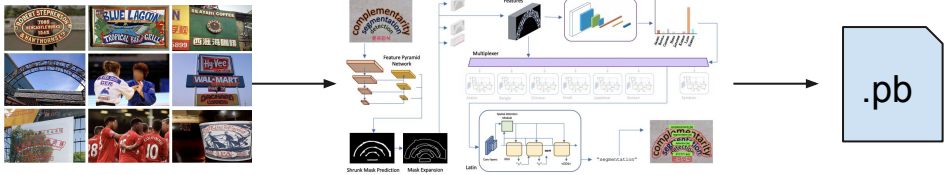
dataset M3 TextSpotter model

One single model to deploy and maintain



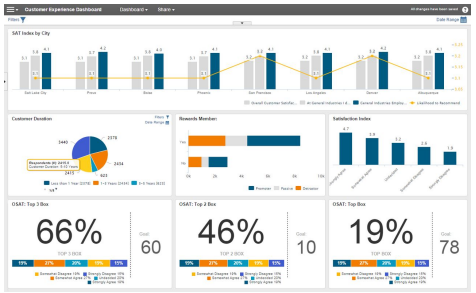
4 Potential value

End-to-End framework that trains in an end-to-end manner



dataset M3 TextSpotter model

One single model to deploy and maintain

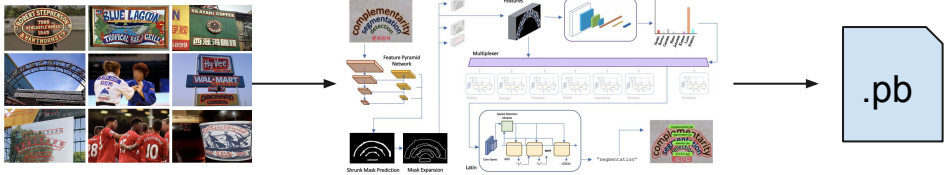


Add, remove or update the recognition module with the languages that you want

Detect and recognize text from multiple languages **within the same scene** → one image

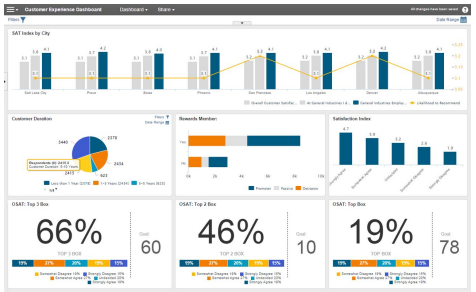
4 Potential value

End-to-End framework that trains in an end-to-end manner



dataset M3 TextSpotter model

One single model to deploy and maintain



Add, remove or update the recognition module with the languages that you want

Detect and recognize text from multiple languages **within the same scene** → one image

Consider before starting:

- Number of languages
- Hardware resources
- Performance goals

veriff

Thank you!

Questions & Answers

Candidate: Víctor Segura Tirado

Paper: A Multiplexed Network for End-to-End, Multilingual OCR