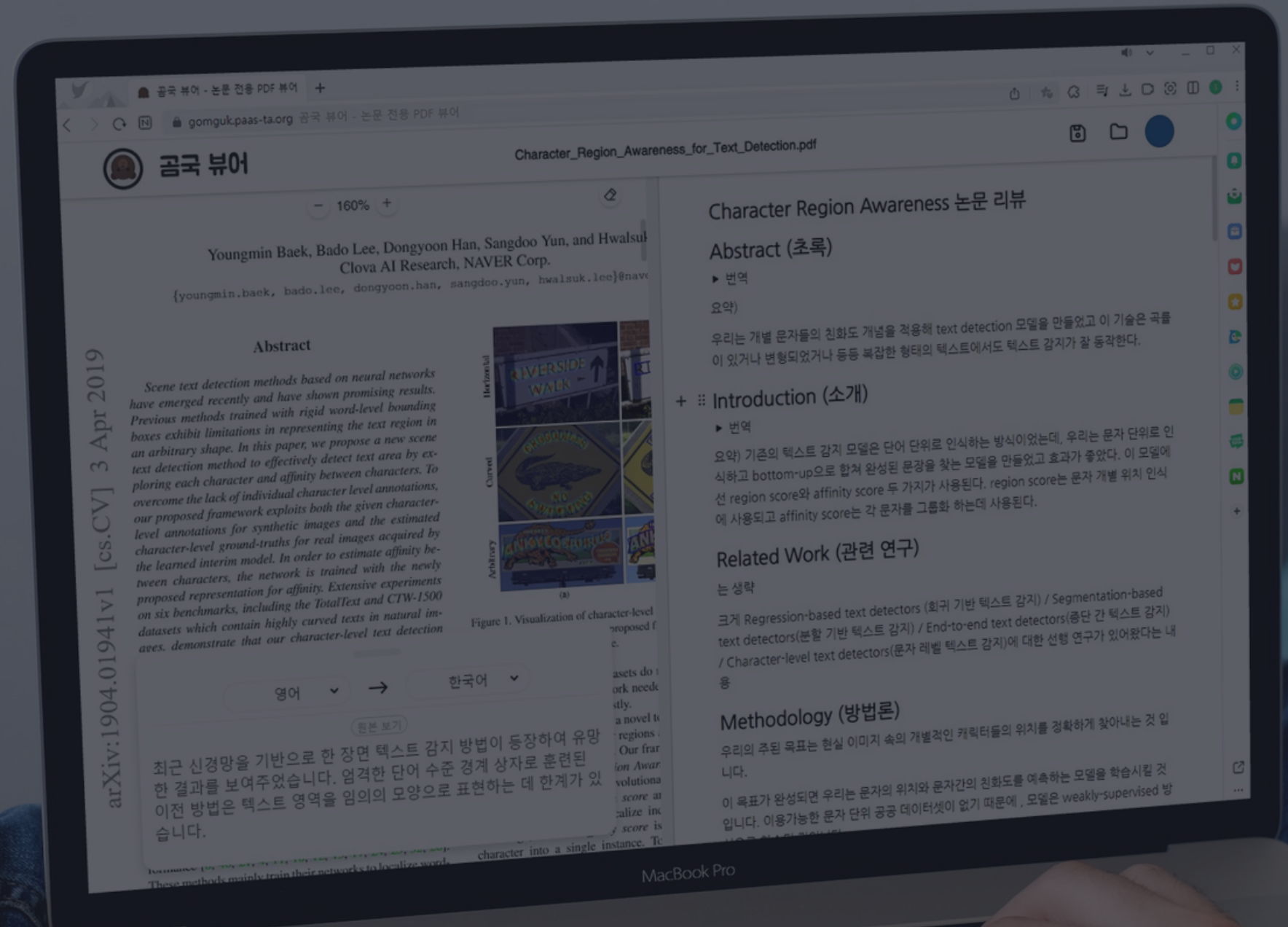




GOMGUK VIEWER
BY 메모리얼

곰국 뷰어

논문 전용 PDF 뷰어



제6 회 개방형 클라우드 플랫폼 기반
서비스 개발 아이디어 공모전

목차

문제 인식

유저 사전 인터뷰

솔루션

서비스 개요

서비스 아키텍처 & 기술 스택

화면 & 기능 설명

서비스 시연

기대효과

추후 계획 & 방향성

영어 해석하기 어렵다..

중요한 내용
정리해야 하는데...

논문 읽기 너무 힘들다

집에서 읽던 논문을
카페에서 이어 읽을 수 없을까?

번역기, 에디터, PDF 뷰어...
에디터를 어디 띄워뒀더라?

논문 파일이
정리가 안되네

유저 사전 인터뷰

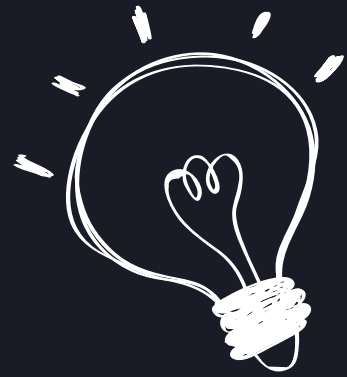


"아이패드로 논문을 주로 읽는데, 한 화면에 3개 이상의 앱을 띄울 수 없어 불편했어요. 논문 띄우고, 번역기 띄우고, 검색용 브라우저도 띄워 두고 왔다 갔다 하면서 사용했거든요."

- 대학생 이XX 님 -

"회사에서 읽을 때도 있고 카페에서 읽을 때도 있기 때문에 사용하는 기기나 환경이 종종 바뀌어요. (불편사항)"

- 회사원 소XX 님 -



이러한 고민들을 한번에 해결해줄 수
있는 **웹앱**을 만들면 어떨까?

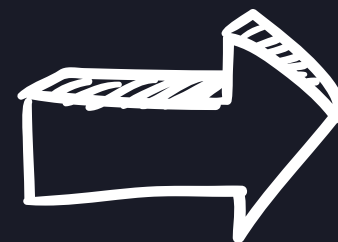
솔루션

필요한 툴들을 이것 저것 띄우다 보니
화면이 너무 복잡해요



번역기, PDF 뷰어, 메모장을 동시에 사
용 가능한 **All-in-One** 형태의 서비스를
제공

데스크탑에서 읽던 논문을 노트북에서
읽고 싶어요



클라우드 기반의 논문 뷰어를 제공해
인터넷만 있으면 어디서든 **동일 환경**의
서비스 이용 가능

GOMGUK VIEWER

솔루션

논문 읽기 경험 개선



GOMGUK VIEWER
서비스 개요

PDF 뷰어



메모장



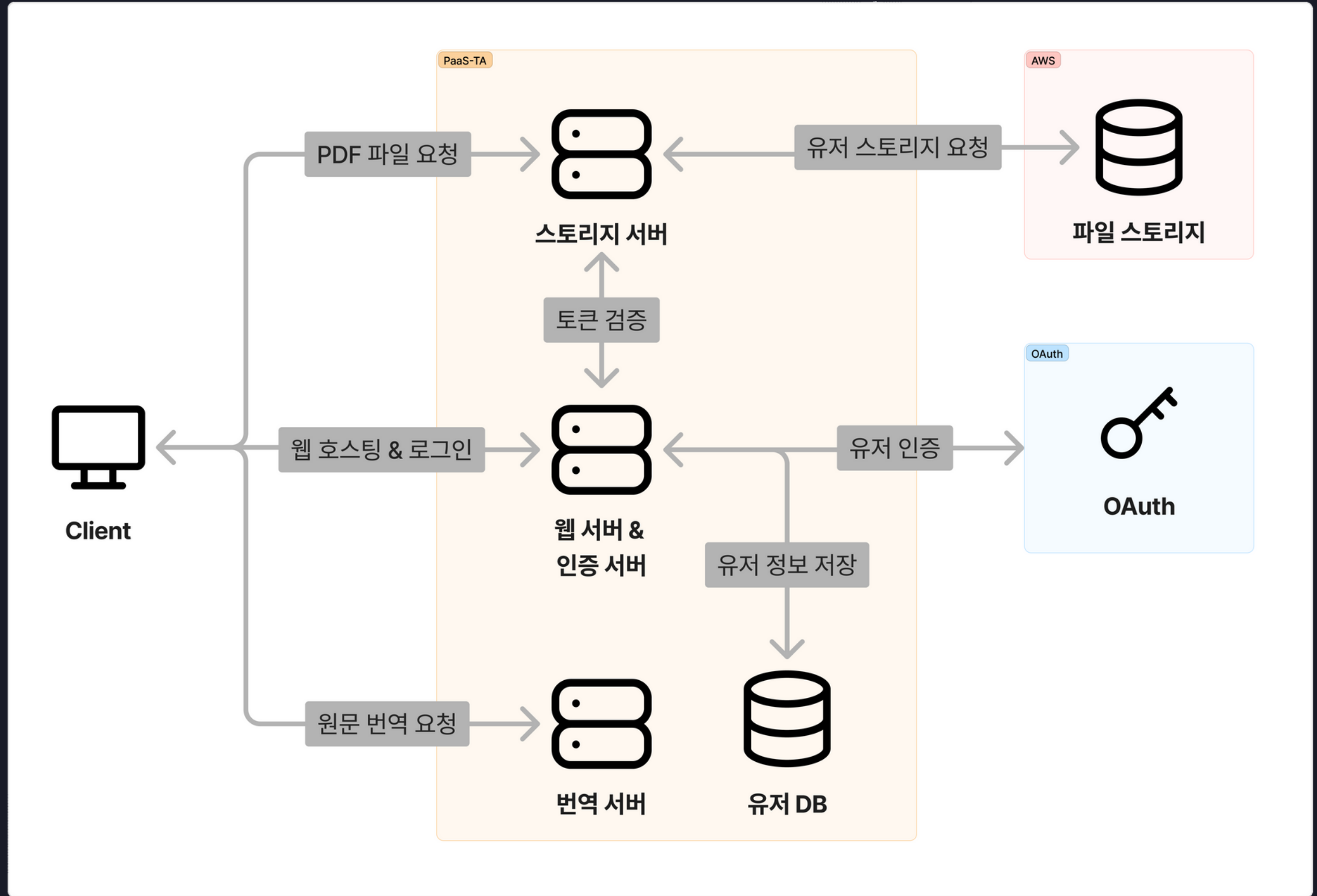
번역



클라우드 스토리지









서비스 아키텍처








기술 스택

Client


 프레임워크 Vue.js	 언어 Typescript
 CSS 전처리기 Scss	 상태관리 Pinia
 에디터 Editor.js	 pdf 뷰어 PDF.js

Cloud Platform



클라우드 플랫폼
PaaS-TA

<p>Translate Server</p>  번역 서버 Django	<p>Storage Server</p>  스토리지 서버 Django
<p>Web Server & Auth Server</p>   웹 서버, 인증 서버 express.js	


GoogleTrans API


번역 API
Google Trans

AWS Cloud


파일 스토리지
AWS S3



Database


데이터베이스
MariaDB

OAuth


OAuth 2.0
Kakao, Google

Deploy

 코드 호스팅 Github	 컨테이너 배포 CloudFoundry
--	---

공국 뷰어
Character_Region_Awareness_for_Text_Det...
📄 📁 👤

160%
🔍

Youngmin Baek, Bado Lee, Dongyoon Han, Sangdoon Yun, and Hwalsuk Lee
Clova AI Research, NAVER Corp.
{youngmin.baek, bado.lee, dongyoon.han, sangdoon.yun, hwalsuk.lee}@naver.com

Abstract

Scene text detection methods based on neural networks have emerged recently and have shown promising results. Previous methods trained with rigid word-level bounding boxes exhibit limitations in representing the text region in an arbitrary shape. In this paper, we propose a new scene text detection method to effectively detect text area by exploring each character and affinity between characters. To overcome the lack of individual character level annotations, our proposed framework exploits both the given character-level annotations for synthetic images and the estimated character-level ground-truths for real images acquired by the learned interim model. In order to estimate affinity between characters, the network is trained with the newly proposed representation for affinity. Extensive experiments on six benchmarks, including the TotalText and CTW-1500 datasets which contain highly curved texts in natural images, demonstrate that our character-level text detection significantly outperforms the state-of-the-art detectors. According to the results, our proposed method guarantees high flexibility in detecting complicated scene text images, such as arbitrarily-oriented, curved, or deformed texts.

1. Introduction

최근 신경망을 기반으로 한 장면 텍스트 감지 방법이 등장하여 유망한 결과를 보여주었습니다.

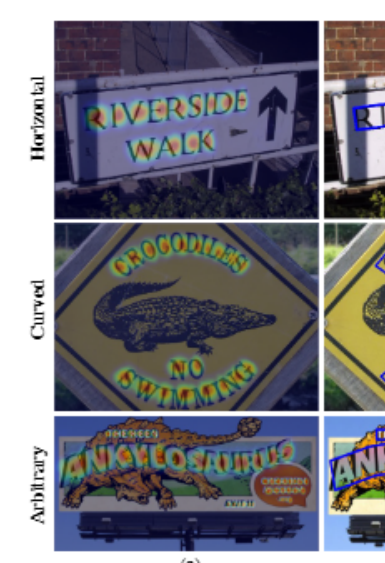


Figure 1. Visualization of character-level detection results for texts of various shape. (a) Heatmaps predicted by our proposed framework for texts of various shape.

most of the existing text datasets do not provide individual character level annotations, and the work needed to generate individual ground truths is too costly. In this paper, we propose a novel text detection framework that utilizes the individual character regions and their affinity. Our framework is aware of the individual character regions and their affinity between characters. This is the first time that character-level awareness is introduced to scene text detection. To overcome the lack of individual character level annotations, our proposed framework exploits both the given character-level annotations for synthetic images and the estimated character-level ground-truths for real images acquired by the learned interim model. In order to estimate affinity between characters, the network is trained with the newly proposed representation for affinity. Extensive experiments on six benchmarks, including the TotalText and CTW-1500 datasets which contain highly curved texts in natural images, demonstrate that our character-level text detection significantly outperforms the state-of-the-art detectors. According to the results, our proposed method guarantees high flexibility in detecting complicated scene text images, such as arbitrarily-oriented, curved, or deformed texts.

영어
→
한국어

[원본 보기](#)

Alternatively, character-level awareness has many advantages. For example, it can be used for a visualization of CRAFT's character-level awareness. This visualization is shown in Figure 1. Figure 1 shows heatmaps predicted by our proposed framework for texts of various shape. (a) Heatmaps predicted by our proposed framework for texts of various shape.

Abstract

이전 메소드들 => 텍스트 영역을 임의의 모양으로 표현하는데 한계가 있다.

+ :: 텍스트 영역을 감지하는데 효과적인 모형 제시 (by 문자 간의 관계(=거리?)를 탐색함으로써)

- 프레임워크가 이용하는 것
 - 합성된 이미지의 문자 어노테이션
 - 학습된 모델로부터 직접 측정된 정보(ground-truth)

확장실험 결과 일상 이미지의 많이 휘어진 텍스트들도 최첨단 텍스트 탐지기를 능가하는 퍼포먼스를 보였다. => 변형, 휘어진, 임의 방향의 텍스트들에 높은 대응성을 가진다.

1. Introduction

최근의 딥러닝 기반 장면 텍스트 탐색기는 유망한 퍼포먼스를 보여준다.

이것들은 올바른 모양의 텍스트에 대해 학습되었을 것이고 그렇기에 휘어진 텍스트들, 극도로 긴 글 등 단일 범위 박스로 감지하기 어려운 텍스트에는 어려움을 겪을 것임. 꼭 그렇지 않더라도 이것들을 다루기 위해 많은 어드밴티지들이 필요할 것

텍스트 인스턴스에 대한 지역화를 해주는 'novel text detector' 소개함 (=CRAFT)

CRAFT는 cnn으로 설계됨

- CRAFT elements
 - region score : 로컬라이징 역할
 - affinity score : 각 문자를 단일 인스턴스로 그룹화하는 역할

문자 어노테이션 부족 보강을 위해 일상 데이터를 사용함으로써 약간 supervise함

2. Related Work

딥러닝 이전 -> 수작업이 주를 이룸(ex. MSER, SWT)

딥러닝 이후 -> SSD, Faster R-CNN 등 유명한 객체탐지기를 채택한 텍스트 감지기가 대부분

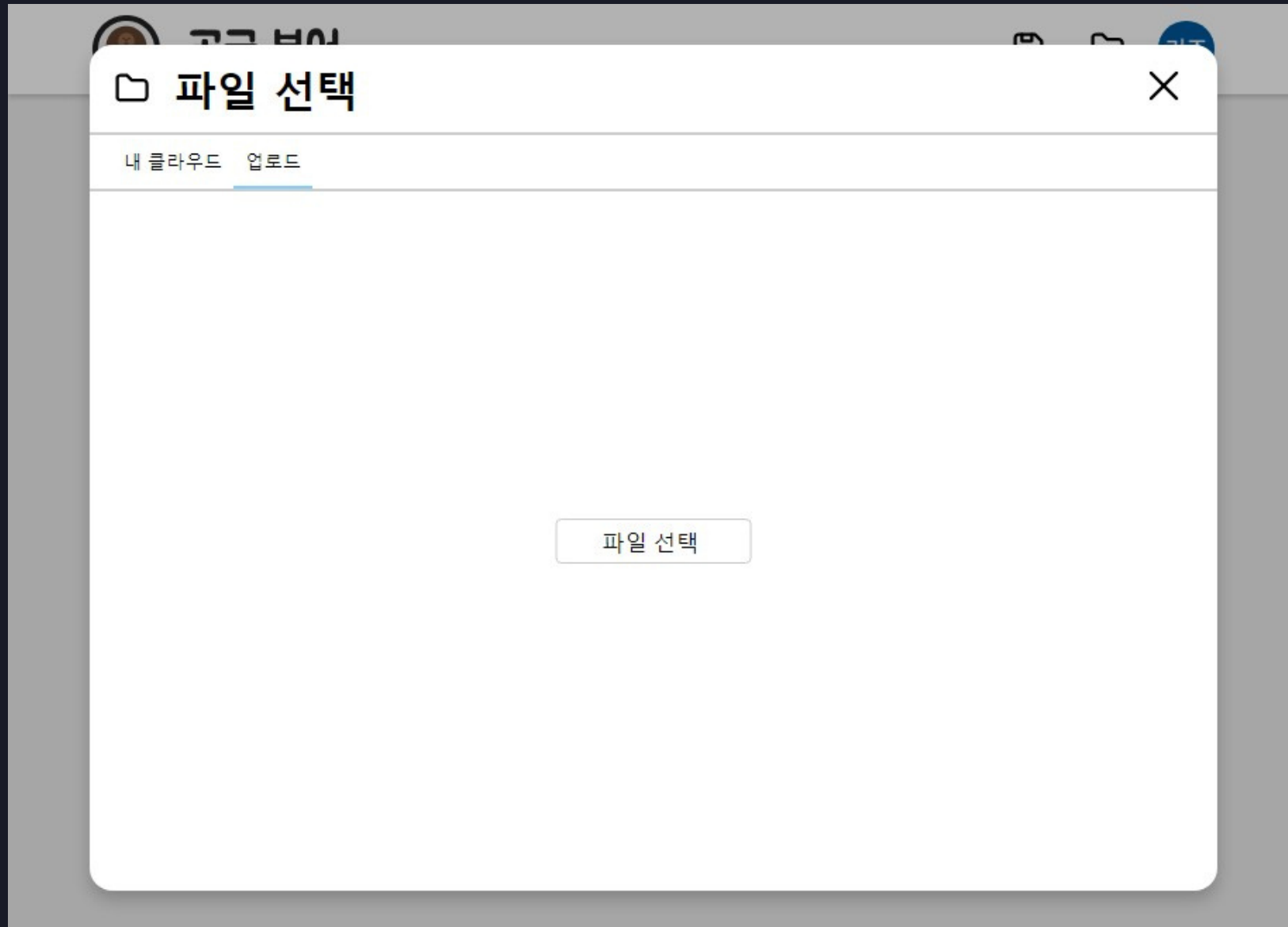
회귀 기반 감지기

일반적인 물체와 다르게 텍스트는 다양한 모양과 각도를 가짐

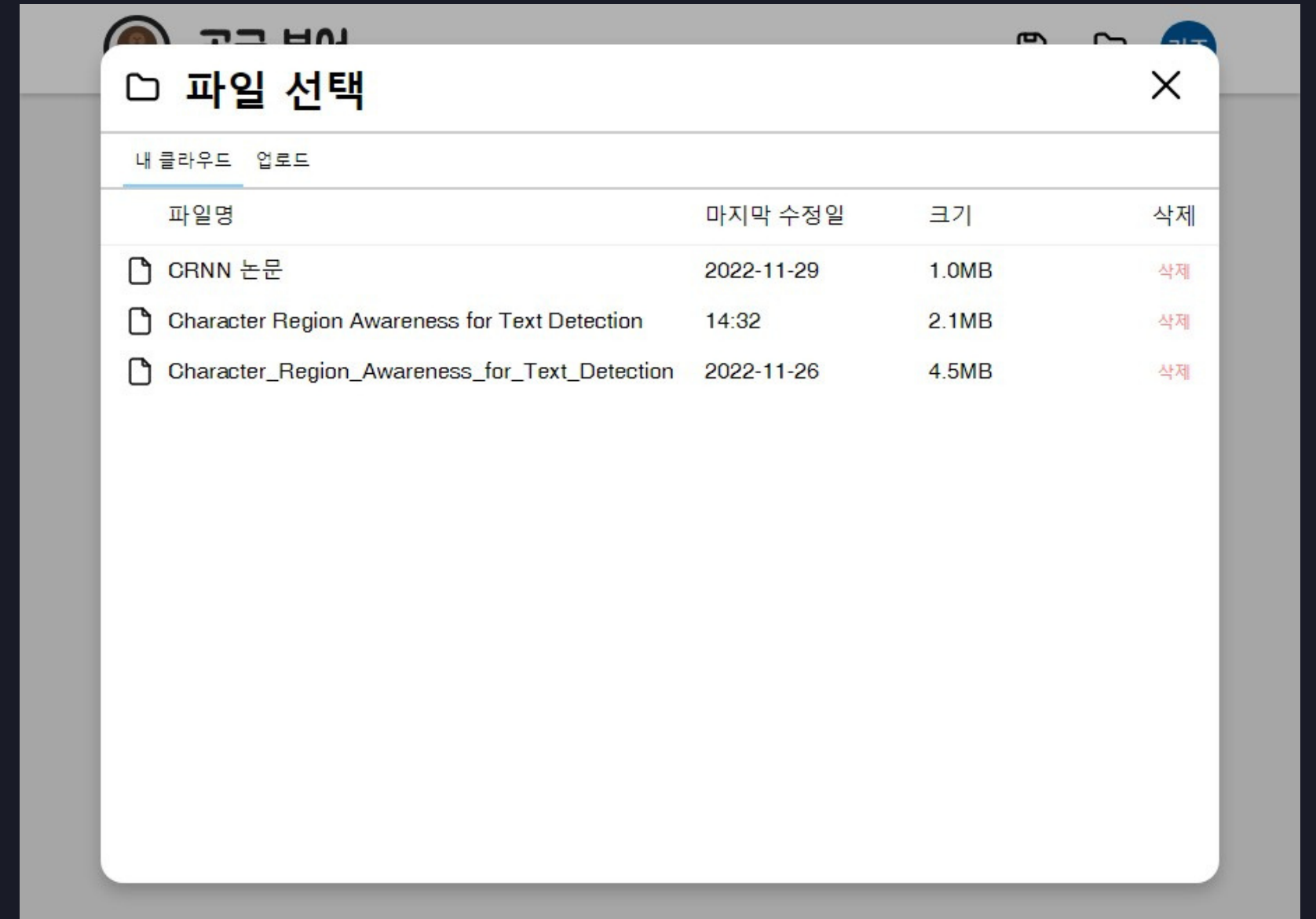
-> 기존 탐지기는 실상에 존재하는 텍스트 모양들을 모두 캐치하기에는 구조적 제한이 있다.

arXiv:1904.01941v1 [cs.CV] 3 Apr 2019

기능 소개 - 클라우드 스토리지



로컬 파일을 클라우드에 업로드



클라우드에서 파일 불러오기



arXiv:1910.04396v1 [cs.CV] 10 Oct 2019

Abstract

Scene text recognition (STR) is the task of recognizing character sequences in natural scenes. While there have been great advances in STR methods, current methods still fail to recognize texts in arbitrary shapes, such as heavily curved or rotated texts, which are abundant in daily life (e.g. restaurant signs, product labels, company logos, etc). This paper introduces a novel architecture to recognizing texts of arbitrary shapes, named *Self-Attention Text Recognition Network (SATRN)*, which is inspired by the Transformer. SATRN utilizes the self-attention mechanism to describe two-dimensional (2D) spatial dependencies of characters in a scene text image. Exploiting the full-graph propagation of self-attention, SATRN can recognize texts with arbitrary arrangements and large inter-character spacing. **As a result, SATRN outperforms existing STR models by a large margin of 5.7 pp on average in “irregular text” benchmarks.** We provide empirical analyses that illustrate the inner mechanisms and the extent to which the model is applicable (e.g. rotated and multi-line text). We will open-source the code.

Introduction

Scene text recognition (STR) addresses the following problem: given an image patch tightly containing text taken from natural scenes (e.g. license plates and posters on the street), what is the sequence of characters? (Zhu, Yao, and Bai 2016; Long, He, and Ya 2018) Applications of deep neural networks have led to great improvements in the performance of STR models (Shi et al. 2016; Lee and Osindero 2016; Yang et al. 2017; Cl



Figure 1: Texts of arbitrary shapes: remaining challenges for scene text recognition.

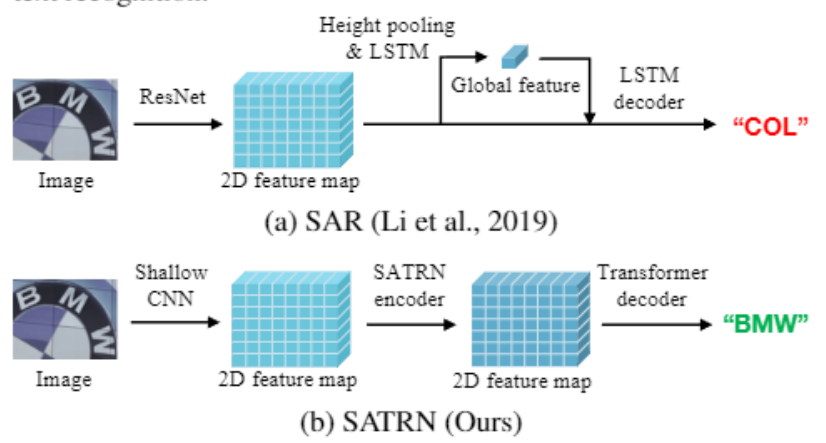


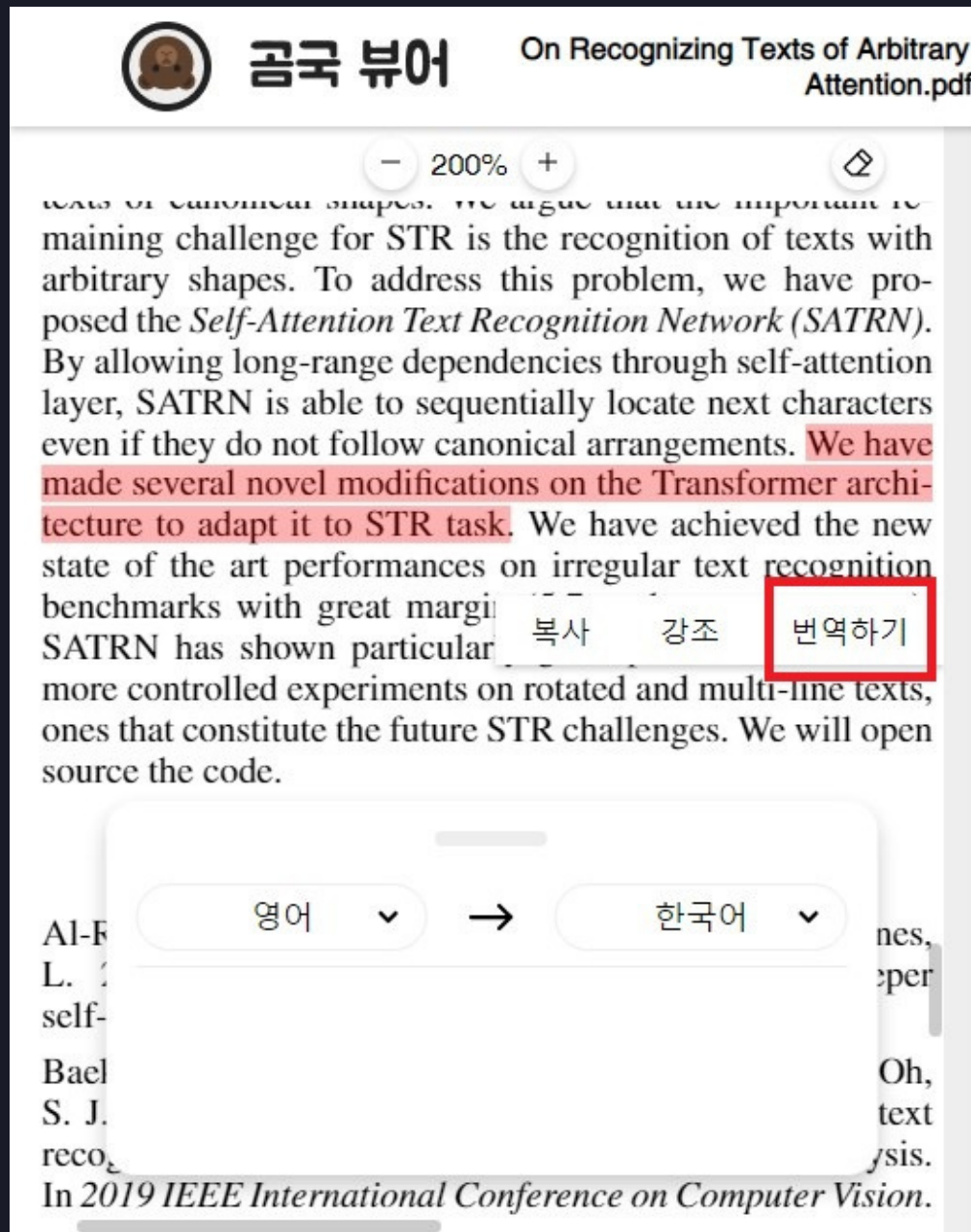
Figure 2: SATRN addresses the text images of difficult shapes (curved “BMW” logo) by adopting a self-attention mechanism, while keeping intermediate feature maps two dimensional. SATRN thus models long-range dependencies spanning 2D space, a feature necessary for recognizing texts of irregular geometry.

interpreting texts with arbitrary shapes, which are important challenges in realistic deployment scenarios.

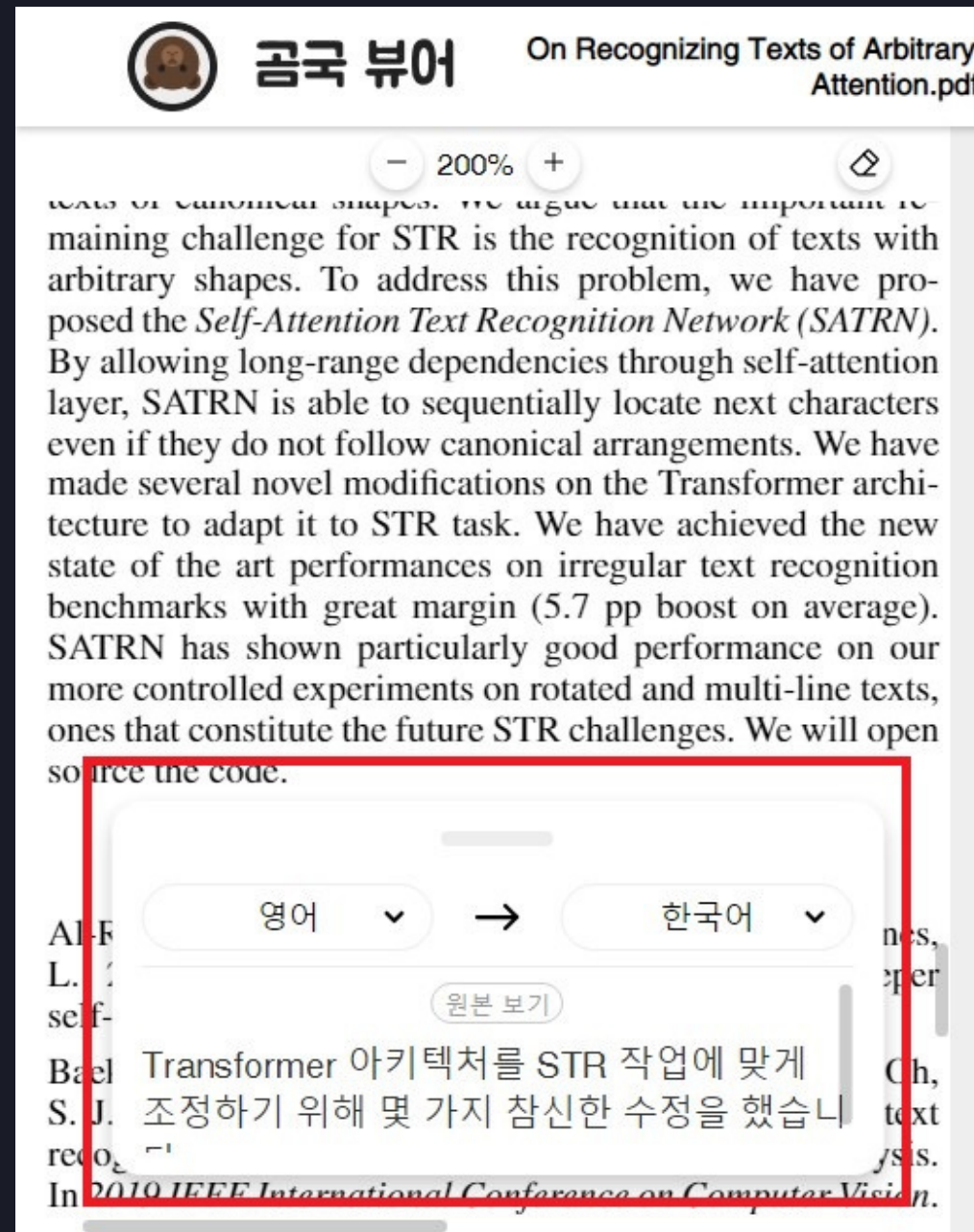
Realizing the significance and difficulty of recognizing as put more emphasis on such image types. The introduction of “irregular shape” STR benchmarks (Back et al. 2019) is an evi-

에 디 터

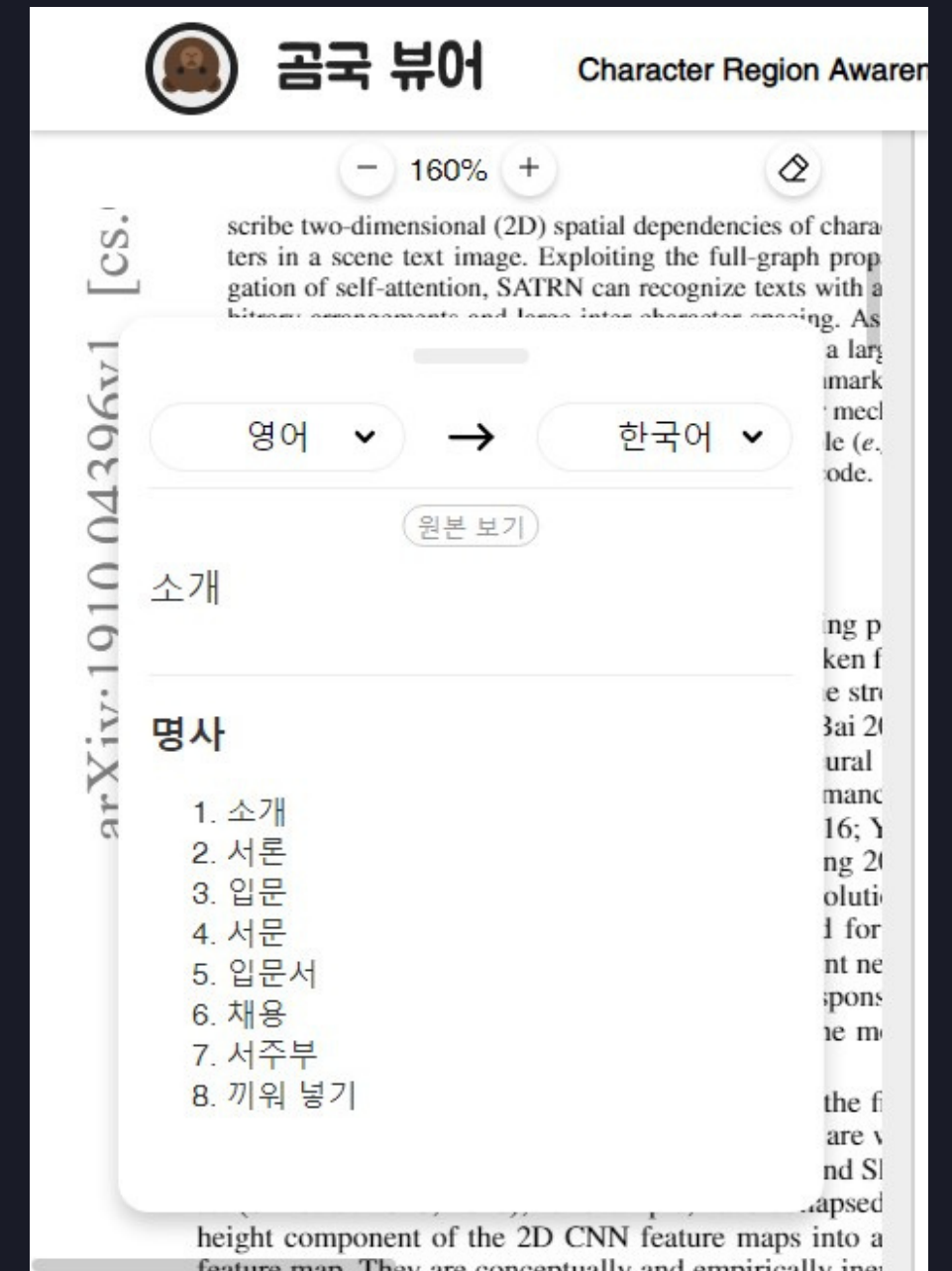
기능 소개 - 번역하기



번역할 텍스트 드래그



원하는 언어에 맞게 번역



단어 번역 시 다양한 뜻

기능 소개 - 메모장

Texts of Arbitrary Shapes with 2D Self-Attention.pdf

2D Self-Attention 논문 리뷰

주요 내용


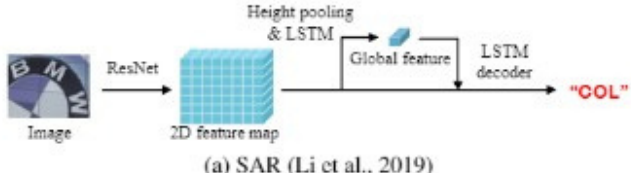
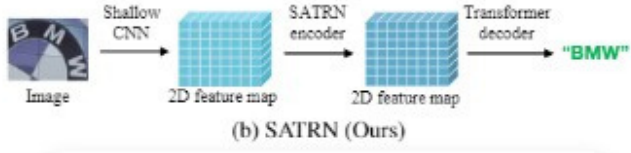


Figure 1: Texts of arbitrary shapes: remaining challenges for scene text recognition.

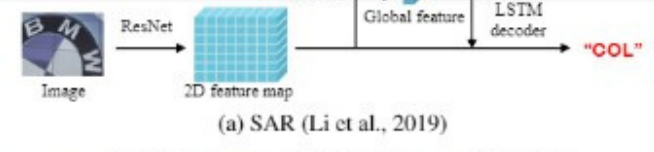
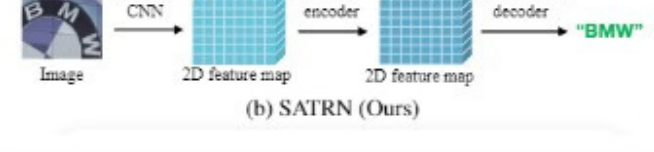



SATRN

임의의 형태의 텍스트 인식의 중요성과 어려움을 깨닫고 STR 커뮤니티는 이러한 이미지 유형에 더 중점을 두었습니다.

기본적인 에디터

Texts of Arbitrary Shapes with 2D Self-Attention.pdf

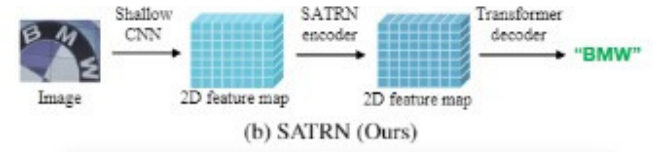
SATRN

임의의 형태의 텍스트 인식의 중요성과 어려움을 깨닫고 STR 커뮤니티는 이러한 이미지 유형에 더 중점을 두었습니다.

- Filter
- T** Text
- H** Heading
- ≡** List
- ▶** Toggle
- α** latex

에디터 블록 종류

Texts of Arbitrary Shapes with 2D Self-Attention.pdf

SATRN

임의의 형태의 텍스트 인식의 중요성과 어려움을 깨닫고 STR 커뮤니티는 이러한 이미지 유형에 더 중점을 두었습니다.

$$P_{p,2i}^{sinu} = \sin(p/10000^{2i/D})$$

라텍스 블록 기능



서비스 시연

<https://gomguk.paas-ta.org/>

기대 효과

논문 읽는데 어려움을 겪는 사용자들의 **피로감 해소**

번역기, PDF 뷰어, 메모장 등을 한 번에
켜놓고 관리하며 논문을 읽어야 했던 기존 경험을
개선하여 피로감을 해소해 줄 수 있습니다

논문들을 한곳에서 관리할 수 있는 **통합 환경 제공**

이 논문은 데스크톱에 저 논문은 노트북에... 정신없이
흩어져 있던 내 논문 파일들을 클라우드를 통해 한곳에
모아두고 관리할 수 있습니다

GOMGUK VIEWER

추후 계획 & 방향성



논문 댓글 기능 추가



태블릿에서의
사용성 증대



연관 논문 추천
시스템 구축

A group of diverse people, including men and women of various ethnicities, are gathered in a meeting room. They are all smiling and shaking hands, suggesting a successful collaboration or agreement. The background shows a wooden wall with a grid of small square openings. The overall atmosphere is positive and professional.

감사합니다