



SCAN ME



Chula
Chulalongkorn University

Practical AI for Text Detection and Recognition in Land Title Deed Documents

Teerapong Panboonyuen

Chulalongkorn University

<https://github.com/kaopanboonyuen/LandDeedOCR>

Reference

- Panboonyuen, Teerapong, Chaiyut Charoenphon, and Chalermchon Satirapod. "CHULA: Custom Heuristic Uncertainty-Guided Loss for Accurate Land Title Deed Segmentation." IEEE Access (2025).
- Wang, Zilong, and Xiaoyu Shen. "Hybrid OCR-LLM Framework for Enterprise-Scale Document Information Extraction Under Copy-heavy Task." arXiv preprint arXiv:2510.10138 (2025).
- Wei, Haoran, Yaofeng Sun, and Yukun Li. "Deepseek-ocr: Contexts optical compression." arXiv preprint arXiv:2510.18234 (2025).
- Memon, Jamshed, et al. "Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR)." IEEE access 8 (2020): 142642-142668.
- Garrido-Munoz, Carlos, Antonio Rios-Vila, and Jorge Calvo-Zaragoza. "Handwritten Text Recognition: A Survey." arXiv preprint arXiv:2502.08417 (2025).
- SCB AI Team. "ThaiOCRBench: A Task-Diverse Benchmark for Vision-Language Understanding in Thai." arXiv preprint arXiv:2511.04479 (2025).
- Cui, Cheng, et al. "Paddleocr 3.0 technical report." arXiv preprint arXiv:2507.05595 (2025).



Disclaimer:

This lecture is for **educational purposes only** to teach **responsible use of AI** for OCR and text recognition. It is **not intended for commercial or illegal use** and should be used strictly for **academic learning**.

Motivation & Context

- Land title deeds are **high-value legal documents**
- Require **high accuracy, traceability, and robustness**
- Manual processing is:
 - Time-consuming
 - Error-prone
 - Hard to scale

Goal: Use **Practical AI** to automate text extraction **safely and reliably**

What is **OCR**?

OCR (Optical Character Recognition) is the process of:

1. Taking an image or scanned document
2. Detecting text regions
3. Converting text images into machine-readable text

OCR = Seeing text + Understanding text

OCR Pipeline (High-Level)

Typical OCR system consists of:

1. **Image Preprocessing**
2. **Text Detection** – *Where is the text?*
3. **Text Recognition** – *What does the text say?*
4. **Post-processing** – correction, formatting, validation

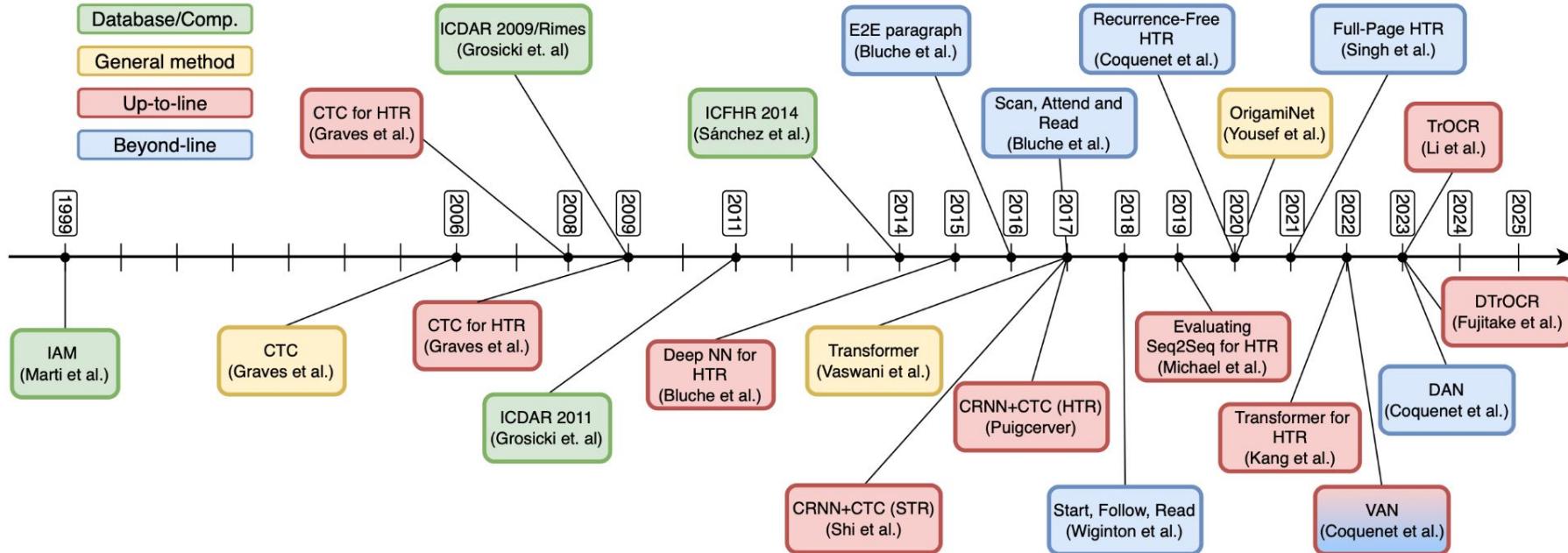


Fig. 2. Timeline of milestones in Handwritten Text Recognition (HTR). We categorize them into four levels: datasets and competitions (green), general methods/architectures (yellow), up-to-line models (red), and beyond-line models (blue).

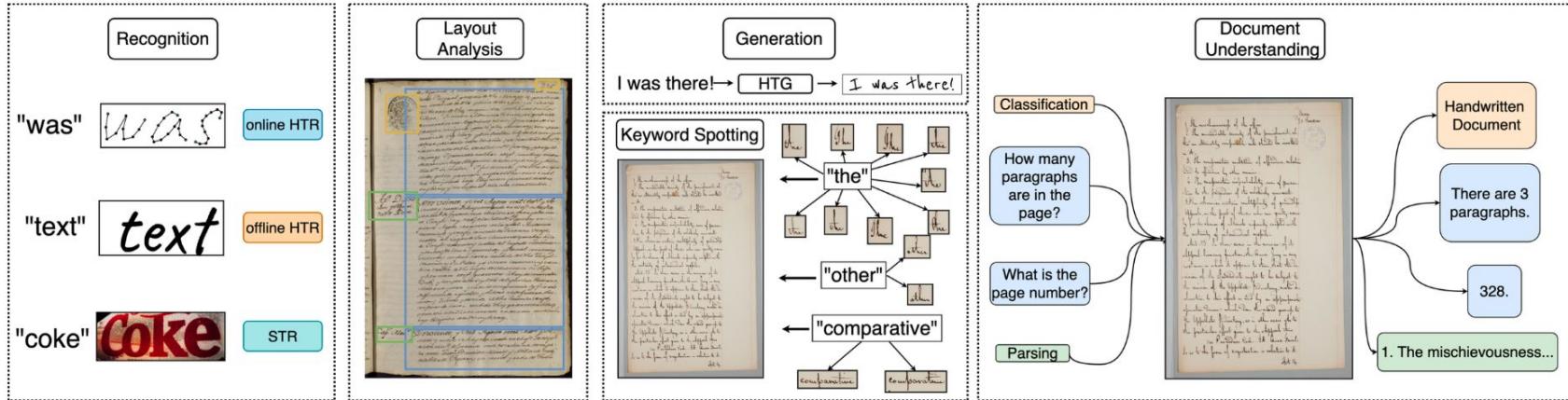


Fig. 6. Overview of related fields in Handwritten Text Recognition (HTR). The figure illustrates key areas closely related to HTR, including: (1) Text Recognition (online HTR, offline HTR, and Scene Text Recognition - STR); (2) Layout Analysis (LA) for segmenting and classifying document regions; (3) Handwritten Text Generation (HTG); (4) Keyword Spotting (KWS) for targeted retrieval of specific words; and (5) Document Understanding (DU), encompassing classification, parsing, and semantic analysis of handwritten documents to extract structured information such as paragraph counts, page numbers, and textual content.

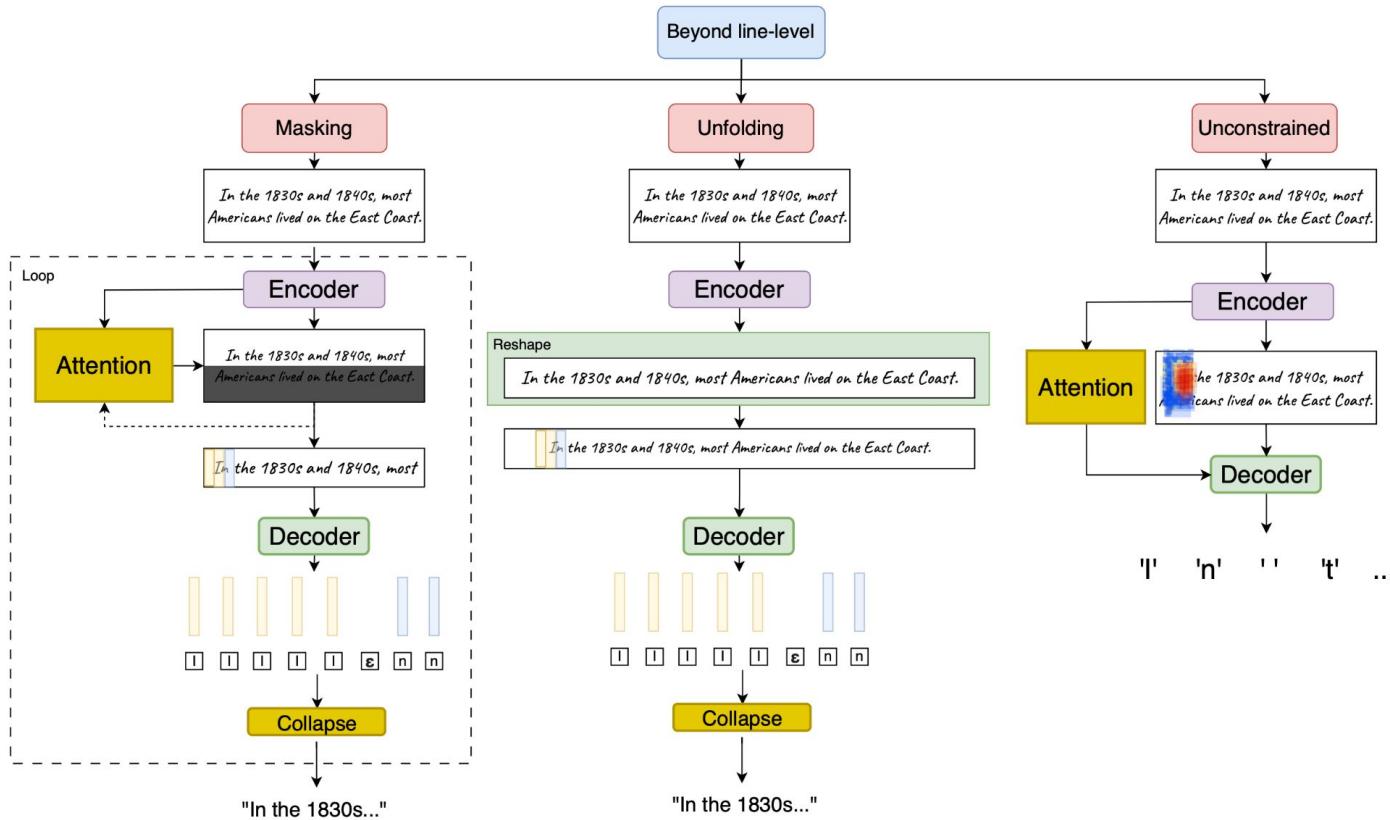
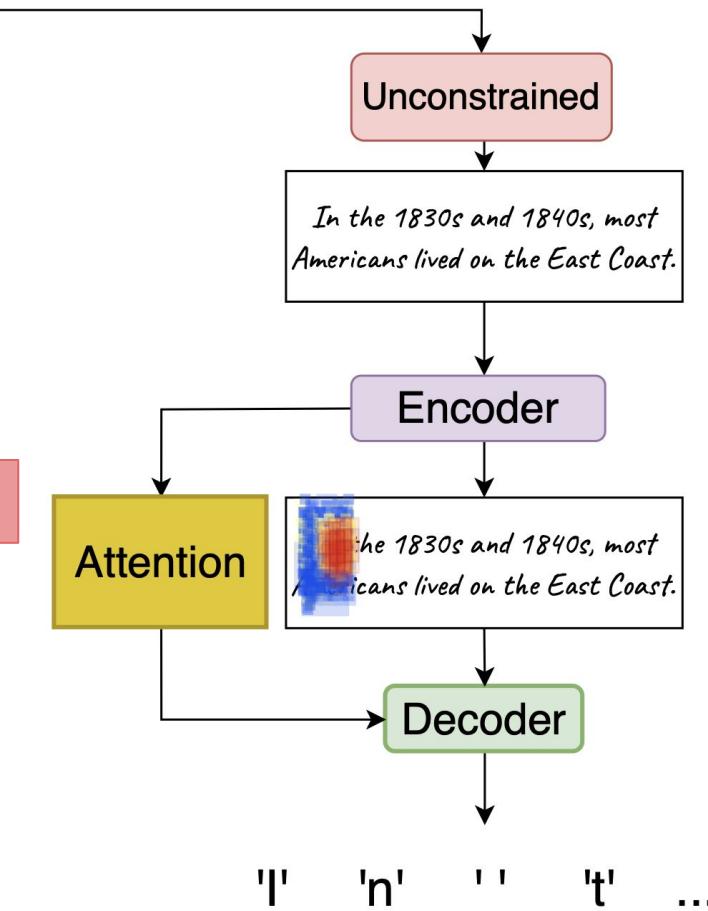
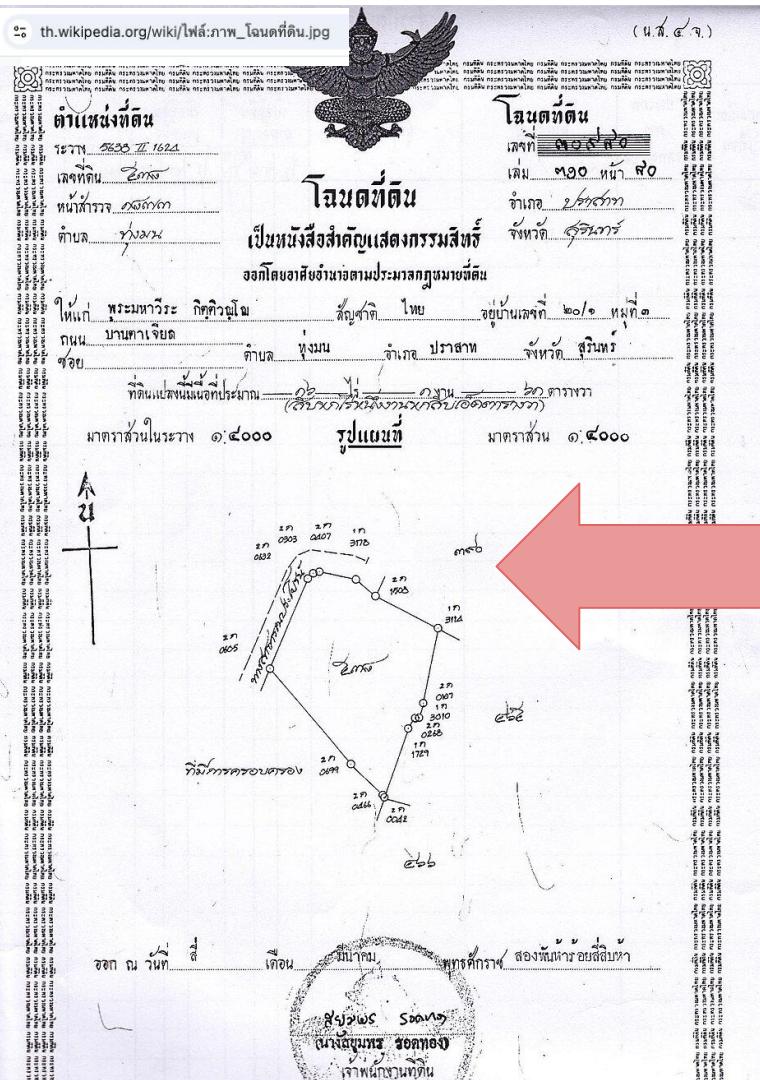
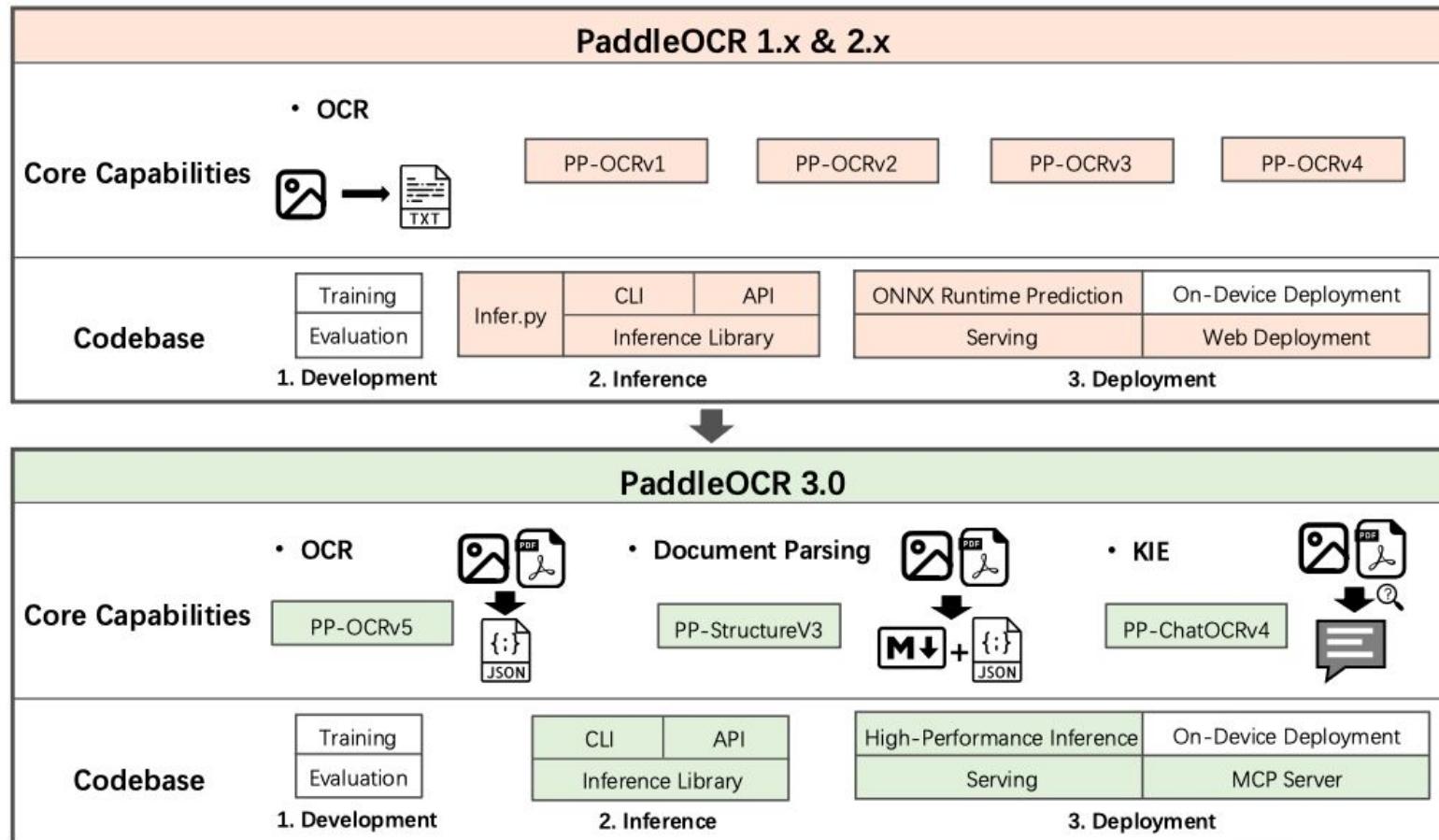


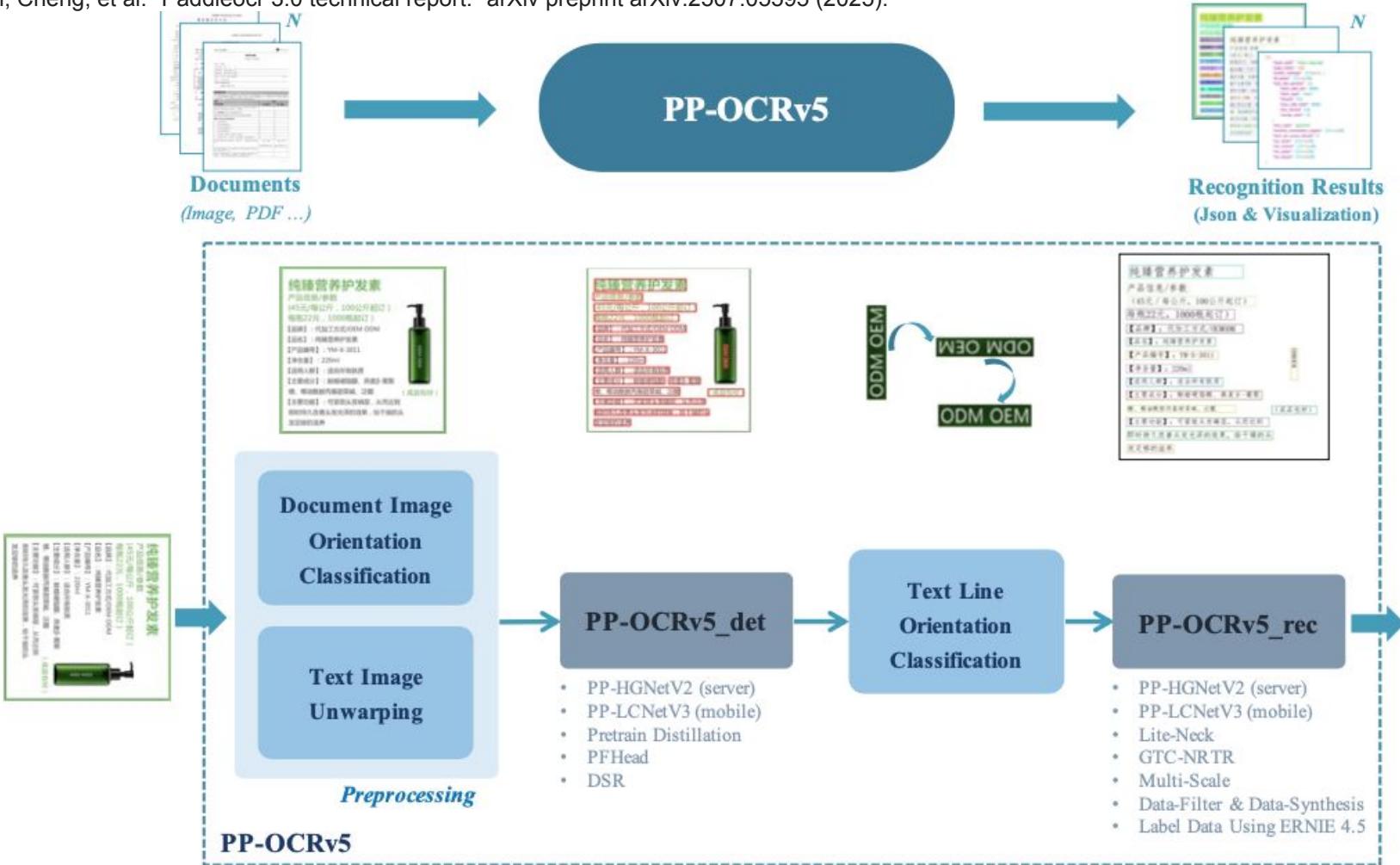
Fig. 11. Taxonomy of HTR approaches beyond line level. Methods are categorized into masking, unfolding, and unconstrained.



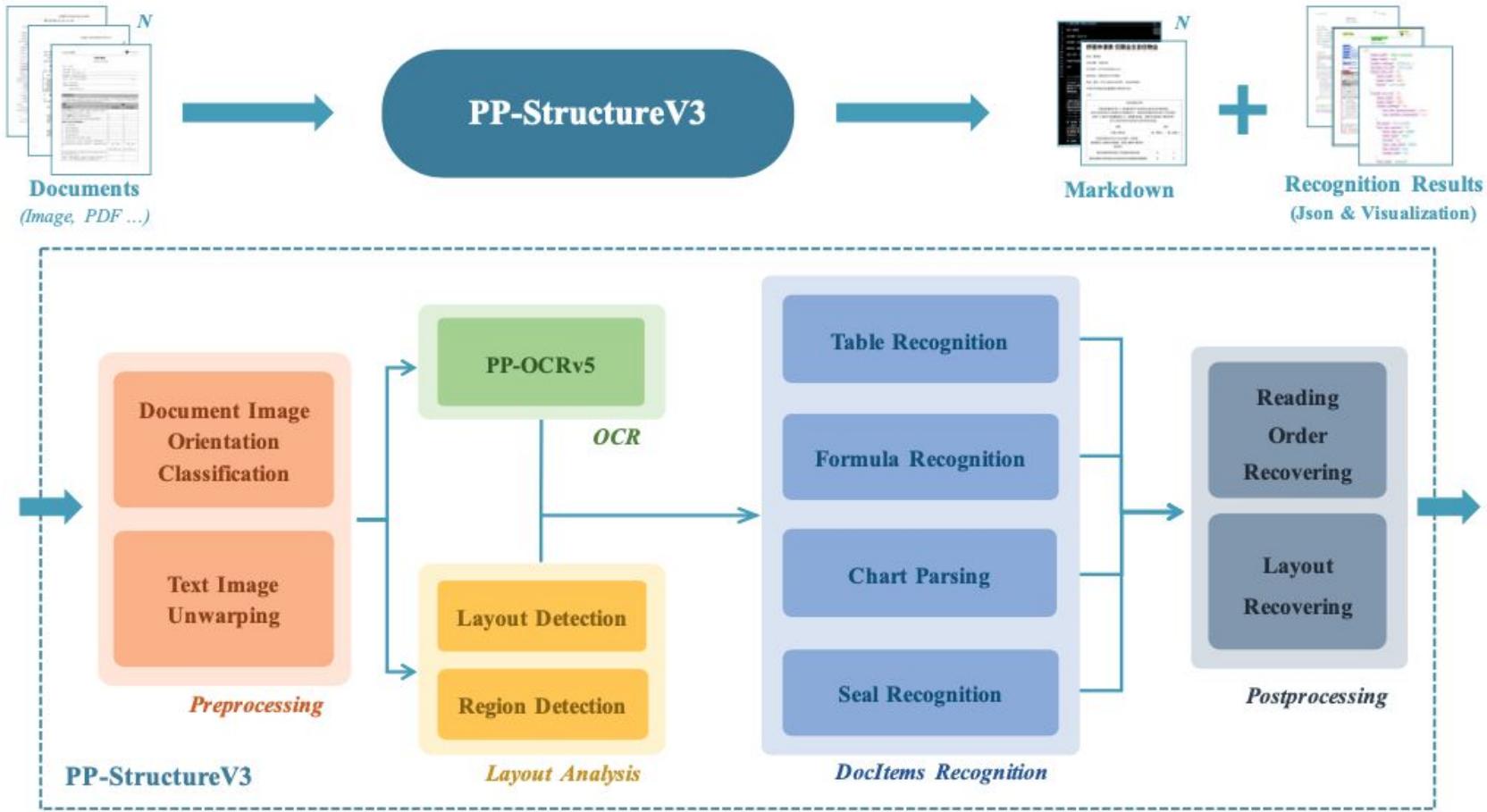
- Cui, Cheng, et al. "Paddleocr 3.0 technical report." arXiv preprint arXiv:2507.05595 (2025).



- Cui, Cheng, et al. "Paddleocr 3.0 technical report." arXiv preprint arXiv:2507.05595 (2025).



- Cui, Cheng, et al. "Paddleocr 3.0 technical report." arXiv preprint arXiv:2507.05595 (2025).



- Cui, Cheng, et al. "Paddleocr 3.0 technical report." arXiv preprint arXiv:2507.05595 (2025).

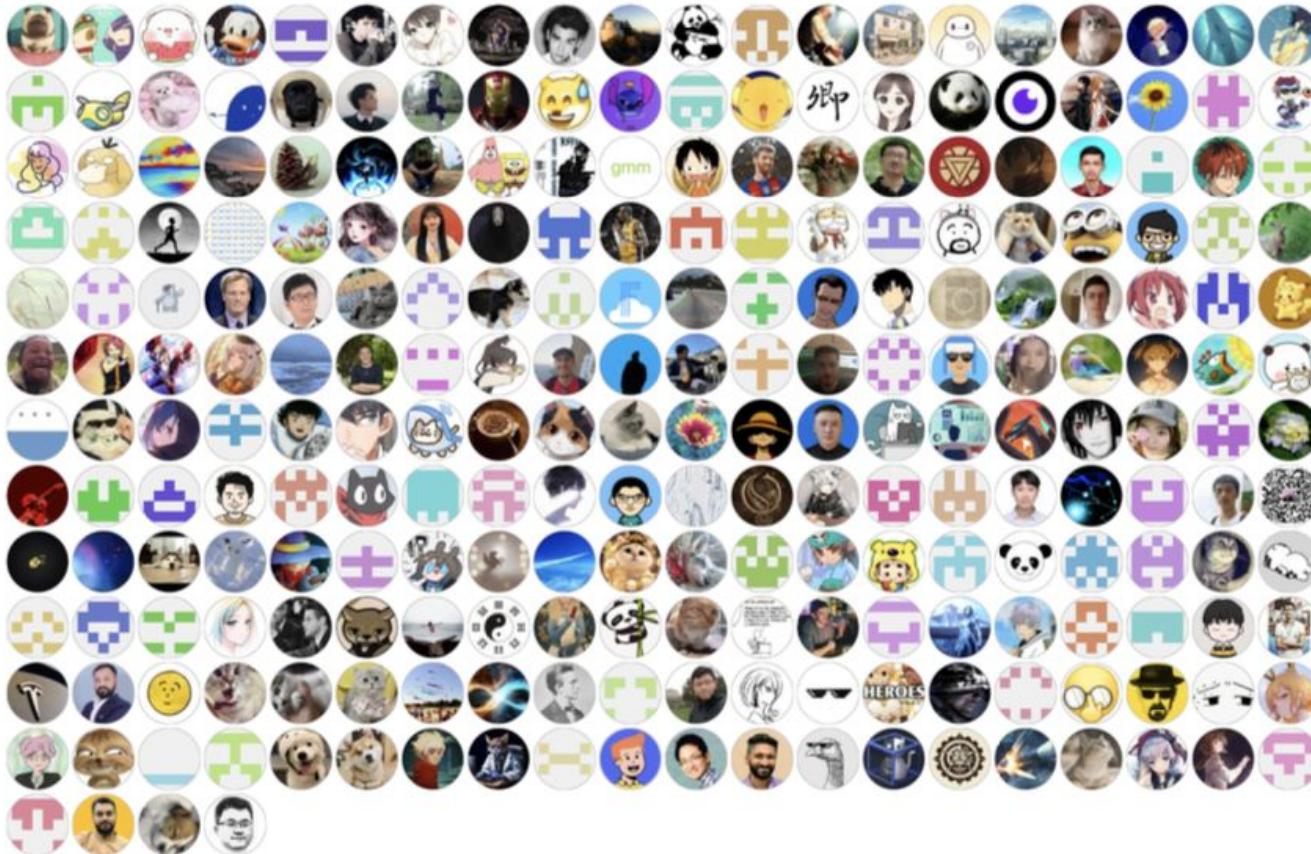


Figure 12 | Avatars of long-term contributors to the PaddleOCR project.

Text Detection vs Text Recognition

Text Detection

- Finds **bounding boxes** of text
- Output: coordinates of text regions

Text Recognition

- Reads characters inside detected boxes
- Output: actual text string

Detection answers *WHERE*, Recognition answers *WHAT*

🎯 Your Mission (Today)

🔍 Step 1 — Text Recognition

Use OCR to extract *all text* from the receipt image

1 2 3 4 Step 2 — Numeric Extraction

Identify and extract **monetary values only**

🧠 Step 3 — AI Reasoning

Handle duplicates, noise, and refunds

💰 Step 4 — Financial Analysis

Compute the **final total cost correctly**



บริษัท ตัวอย่าง จำกัด

2571/1 ช.รามคำแหง 75 ถ.รามคำแหง แขวงหัวหมาก เขตหนองแขม กรุงเทพฯ 10240
โทร. 0-2739-5900 โทรสาร 0-2739-5910 เลขประจำตัวผู้เสียภาษี 31523611000

ใบเสร็จก่อนรับเงิน

รหัสเอกสารนี้: R-0002

ชื่อผู้พนักงาน: บริษัท กรุงเทพมหานครจำกัด (มหาชน)

ที่อยู่:

83/1 หมู่ 4 ถนนพหลโยธิน
ต.คลอง 8 แขวงน้ำดอง กรุงเทพมหานคร 54000
โทร. 035-214-940-2 โทรสาร 035-271-556

เลขที่: RR5311-00001

วันที่: 25/07/2553

ลำดับที่	รายการ	จำนวนเงิน	ยอดคงทิ้ง	ยอดชำระ
1	ใช้รับเงินตามเอกสารเลขที่ DN5302-00001 ลงวันที่ 28/02/2553	2,814,966.70	2,814,966.70	2,814,966.70
2	ใช้รับเงินตามเอกสารเลขที่ BL5302-00001 ลงวันที่ 20/02/2553	3,835,431.05	3,835,431.05	3,835,431.05
3	ใช้รับเงินตามเอกสารเลขที่ ARIN5311-00002 ลงวันที่ 01/01/2553	100,000.00	100,000.00	100,000.00
4	ใช้รับเงินตามเอกสารเลขที่ SN5311-00001 ลงวันที่ 20/02/2553	149,526.08	149,526.08	149,526.08
5	ใช้รับเงินตามเอกสารเลขที่ CN5302-00001 ลงวันที่ 25/02/2553	-97,865.84	-97,865.84	-97,865.84

หมายเหตุ

💰 FINAL RECEIPT TOTAL: 6,802,057.99 THB

That's it! You can go home now 😊

Text Detection

Text Detection: How It Works

Detection models treat text as:

- Objects (like cars or people)
- Or sequences of pixels

Common approaches:

- CNN-based feature extraction
- Geometry or segmentation-based outputs

Popular Text Detection Models (SOTA)

- **EAST** — Efficient and fast
- **DBNet / DBNet++** — Differentiable binarization (widely used)
- **CRAFT** — Character-level detection
- **YOLO-based Text Detectors** — High speed, modern pipelines

DBNet is commonly used in industrial OCR systems

But

CHULA is not a separate detection model — it's a *training objective* that **improves existing**

◆ ຂໍ້ມູນກາພຽມໂດຍ AI

"CHULA title deed" refers to an academic method (Custom Heuristic Uncertainty-guided Loss) developed by Chulalongkorn University (Chula) researchers for automatically segmenting (digitizing) land boundaries from physical Thai land title deeds, improving accuracy for digital land records, rather than a specific *type* of official deed itself. This technique uses AI/machine learning to overcome challenges with old or unclear deeds, making land data more reliable for official purposes like digital transformation of the land registry. ☀

What it is:

- CHULA (Custom Heuristic Uncertainty-guided Loss): A specific algorithm.

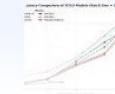
ແສດຖະເກີດເຕີມ ▾

CHULA: Custom Heuristic Uncertainty-guided Loss for Accurate Land Title Deed Segmentation | IEEE...

1 ກ.ຍ. 2568 — CHULA (Custom Heuristic Uncertainty-guided Loss) is a method for segmenting land boundaries from Thai...

IEEE Xplore :

(PDF) CHULA: Custom Heuristic Uncertainty-guided Loss for ...



10 ກ.ຍ. 2568 — CHULA: Custom Heuristic Uncertainty-guided Loss for Accurate Land Tit...

IEEE Xplore

<https://ieeexplore.ieee.org> · ແລ້ວກຳນົດ :

CHULA: Custom Heuristic Uncertainty-guided Loss for ...

ໂດຍ T Panboonyuen · 2025 — Evaluated on a carefully curated Thai Land Title Deed Dataset, CHULA achieves an impressive 92.4% accuracy, significantly surpassing standard ...

ResearchGate

<https://www.researchgate.net> · 395224... · ແລ້ວກຳນົດ :

(PDF) CHULA: Custom Heuristic Uncertainty-guided Loss ...

11 ກ.ຍ. 2568 — Evaluated on a carefully curated Thai Land Title Deed Dataset, CHULA achieves an impressive 92.4% accuracy, significantly surpassing standard ...

← → ⌂ Guest

ieeexplore.ieee.org/document/11146788

IEEE.org IEEE Xplore IEEE SA IEEE Spectrum More Sites Subscribe Donate Cart Create Account Personal Sign In

IEEE Xplore® Browse My Settings Help Institutional Sign In

All ADVANCED SEARCH

Journals & Magazines > IEEE Access > Volume: 13 ?

CHULA: Custom Heuristic Uncertainty-Guided Loss for Accurate Land Title Deed Segmentation

Publisher: IEEE [Cite This](#) [PDF](#)

Teerapong Panboonyuen ; Chaiyut Charoenphon ; Chalermchon Satirapod All Authors

200 Full Text Views

Open Access Comment(s)

Under a Creative Commons License

Abstract

Abstract: Accurately segmenting land boundaries from Thai land title deeds is crucial for reliable land management and legal processes, but remains challenging due to low-quality scans, diverse layouts, and complex overlapping elements in documents. Existing methods often struggle with these difficulties, resulting in imprecise delineations that can cause disputes or inefficiencies. To address these issues, we propose CHULA, a novel Custom Heuristic Uncertainty-guided Loss tailored specifically for robust land title deed segmentation. CHULA uniquely combines domain-specific heuristic priors with uncertainty modeling in a unified loss function that effectively guides the model to focus on clearer regions while refining boundaries and suppressing noisy

Document Sections I. Introduction II. Related Work III. Approach

More Like This

SparseFormer: A Credible Dual-CNN Expert-Guided Transformer for Remote Sensing Image Segmentation With Sparse Point Annotation
IEEE Transactions on Geoscience and Remote Sensing
Published: 2025

Improving Accuracy in Medical Imaging: A YOLO-NAS and SAM

PDF Help Feedback

Received 16 August 2025, accepted 29 August 2025, date of publication 2 September 2025, date of current version 9 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3605218



RESEARCH ARTICLE

CHULA: Custom Heuristic Uncertainty-Guided Loss for Accurate Land Title Deed Segmentation

TEERAPONG PANBOONYUEN[✉], CHAIYUT CHAROENPHON[✉],
AND CHALERMCHEON SATIRAPOD[✉]

Mapping and Positioning from Space (MAPS) Technology Research Center, Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University, Pathum Wan, Bangkok 10330, Thailand

Corresponding author: Chalermcheon Satirapod (chalermchon.s@chula.ac.th)

This research project is supported by the Second Century Fund (C2F), Chulalongkorn University.

ABSTRACT Accurately segmenting land boundaries from Thai land title deeds is crucial for reliable land management and legal processes, but remains challenging due to low-quality scans, diverse layouts, and complex overlapping elements in documents. Existing methods often struggle with these difficulties, resulting in imprecise delineations that can cause disputes or inefficiencies. To address these issues, we propose **CHULA**, a novel Custom Heuristic Uncertainty-guided Loss tailored specifically for robust land title deed segmentation. CHULA uniquely combines domain-specific heuristic priors with uncertainty modeling in a unified loss function that effectively guides the model to focus on clearer regions while refining boundaries and suppressing noisy areas. Evaluated on a carefully curated Thai Land Title Deed Dataset, CHULA achieves an impressive **92.4%** accuracy, significantly surpassing standard segmentation baselines. Our results highlight the promise of integrating uncertainty and heuristic knowledge to enhance segmentation accuracy in complex, real-world documents. The code is publicly available at <https://github.com/kaopanboonyuen/CHULA>

INDEX TERMS Document segmentation, uncertainty-aware learning, custom loss function, land title deed, Thai document analysis.

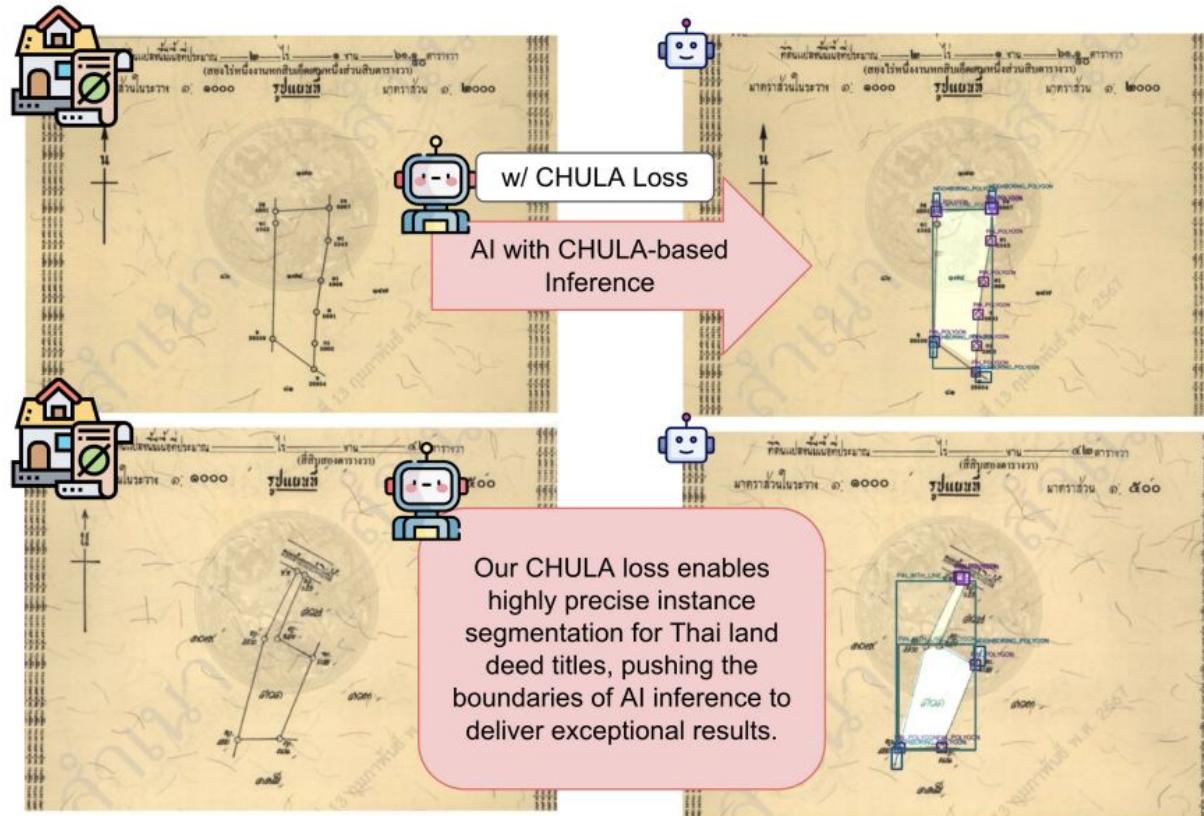


FIGURE 4. Results from our YOLOv12-X-CHULA (Ours). The figure presents two sample test cases in rows 1 and 2. For each row, the left column displays the input image, while the right column shows the predicted segmentation results produced by the model.

TABLE 3. Model efficiency and segmentation performance of CHULA loss variants. FLOPs in Giga-operations; Params in millions; mIoU: mean Intersection over Union.

Loss Variant	FLOPs (G)	Param. (M)	mIoU (%)
Baseline (CE)	91.3	27.4	68.9
CHULA w/o Unc.	92.0	27.4	71.2
CHULA w/o CB	92.2	27.4	72.1
CHULA (Full)	92.2	27.4	73.8

TABLE 5. Comparison of CHULA-integrated models with non-YOLO state-of-the-art segmentation frameworks. CHULA improves performance across all architectures.

Model	$AP_{50:95}^{val}$ (%)	Accuracy (%)	Backbone / Framework
U-Net [34]	53.5	88.7	UNet (ResNet34)
U-Net + CHULA (Ours)	55.6	90.2	UNet (ResNet34)
Mask R-CNN [35]	56.8	89.3	ResNet50-FPN
Mask R-CNN + CHULA (Ours)	58.7	90.8	ResNet50-FPN
DETR [36]	57.4	89.8	Transformer Encoder-Decoder
DETR + CHULA (Ours)	59.6	91.3	Transformer Encoder-Decoder
Swin Transformer [37]	60.3	91.1	Swin-L
Swin Transformer + CHULA (Ours)	62.0	92.0	Swin-L
SegFormer-B4 [38]	59.1	90.5	MiT-B4
SegFormer-B4 + CHULA (Ours)	60.8	91.7	MiT-B4
SAM (ViT-H) [39]	61.2	91.8	Vision Transformer H
SAM + CHULA (Ours)	62.5	92.7	Vision Transformer H

Decision

 Auto-Approved

 Human Review

 Hard Case (Land Title Risk)

DEMO: CHULA AI Performance

New Features:

-  **Auto-Approval:** High-confidence AI can approve without human help.
- **Accuracy Boost:** +2% on tough land deed markers (pins & lines).
- **Smart & Reliable:** Handles complex layouts better than standard OCR.

CHULA = Faster, Smarter, More Accurate

Text Recognition

Text Recognition: Core Idea

Text Recognition converts:

Image of text → Sequence of characters

Challenges:

- Variable length text
- Different fonts and handwriting
- Noise and distortion

Traditional Text Recognition (CRNN)

CRNN = CNN + RNN + CTC

Pipeline:

1. CNN extracts visual features
2. RNN models character sequence
3. CTC aligns prediction without character-level labels

* Convolutional Recurrent Neural Network (CRNN)

CRNN Architecture (Conceptual)

- Input: cropped text image
- CNN → feature map
- RNN → sequence modeling
- CTC loss → final text

Advantages:

- Simple
- Efficient

Limitations:

- Struggles with long or complex layouts

Attention-Based Recognition Models

Modern recognition models use:

- **Encoder–Decoder with Attention**
- **Transformer-based architectures**

Benefits:

- Better long-text handling
- Flexible alignment

SOTA Text Recognition Models

ABINet — Visual + language modeling

PARSeq — Pure transformer-based

TrOCR — Vision–Language pretraining

SVTR — Fast transformer OCR

Why Language Model Matters in OCR

OCR is not only vision, but **language understanding**

Example:

- Visual: "กร**บ**ทีดิน"
- Language-aware OCR: "กร**ม**ทีดิน"

Language models help correct:

- Typos
- Missing characters
- Ambiguous shapes

OCR Evaluation Metrics (Overview)

Common metrics:

- **Character Error Rate (CER)**
- **Word Error Rate (WER)**
- **Accuracy**

Each metric answers a different question

Character Error Rate (CER)

$$\text{CER} = (\text{Sub} + \text{Del} + \text{Ins}) / \text{Total Characters}$$

Example:

- Ground Truth: กรมที่ดิน
- Prediction: กรมท'din

Errors:

- Substitution: 

$$\text{CER} = 1 / 8 = 12.5\%$$

Word Error Rate (WER)

$$\text{WER} = (\text{S} + \text{D} + \text{I}) / \text{Number of Words}$$

Thai challenge:

- No spaces between words
- Requires **word segmentation first**

WER is less reliable for Thai unless preprocessing is applied

Accuracy (Simple View)

Accuracy = Correct Predictions / Total Samples

- Easy to understand
- But hides partial errors

CER is usually preferred for Thai OCR

CER Check 😊

Ground Truth:

เรารักกรมที่ดิน

Prediction:

เรารักกมที่ดิน

👉 Question:

- How many **Sub / Del / Ins** errors?
- What is the **total number of characters**?
- **CER = ?**



Hint: Count character by character — OCR has no mercy.

WER Check 😊

Ground Truth:

เรา รัก กรมที่ดิน

Prediction:

เรา รัก ที่ดิน

👉 Question:

- What type of error occurred?
- What is the **total number of words**?
- **WER = ?**

😊 Hint: Count **words**, not characters.

Why Thai Language is Difficult for OCR

Challenges:

- No explicit word spacing
- Vowels and tones above/below characters
- Similar shapes (ດ / ດ / ຕ)
- Historical fonts in land deeds

Thai OCR requires **language-specific modeling**

Traditional OCR Tools

Tesseract OCR

- Open-source
- Supports Thai
- Rule-based + LSTM

EasyOCR

- Deep learning based
- Fast setup

PaddleOCR

- Industry-grade
- Strong Thai support
- Detection + Recognition + Post-process

Practical Comparison (Tools)

Tool	Accuracy	Customization	Thai Support
Tesseract	Medium	Low	Medium
EasyOCR	Medium-High	Medium	Medium
PaddleOCR	High	High	High

Modern OCR with Multimodal LLMs

New paradigm:

Image + Prompt → Structured Text

Examples:

- OpenAI Vision Models
- Gemini Vision
- Qwen-VL

Strengths:

- Strong layout understanding
- Few-shot learning

OCR with LLMs: Strengths & Risks

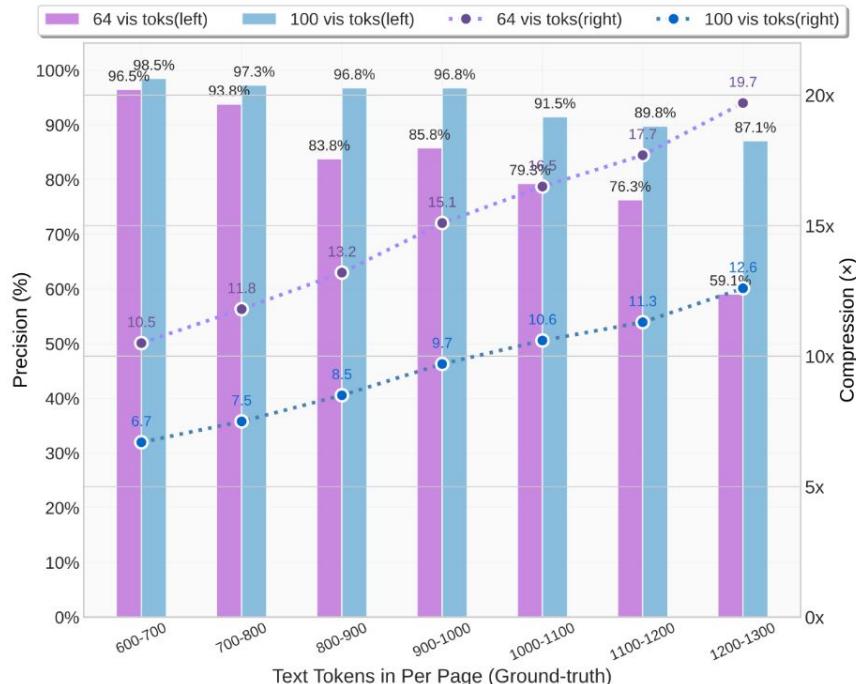
Strengths:

- Flexible
- Good at noisy documents
- Understands context

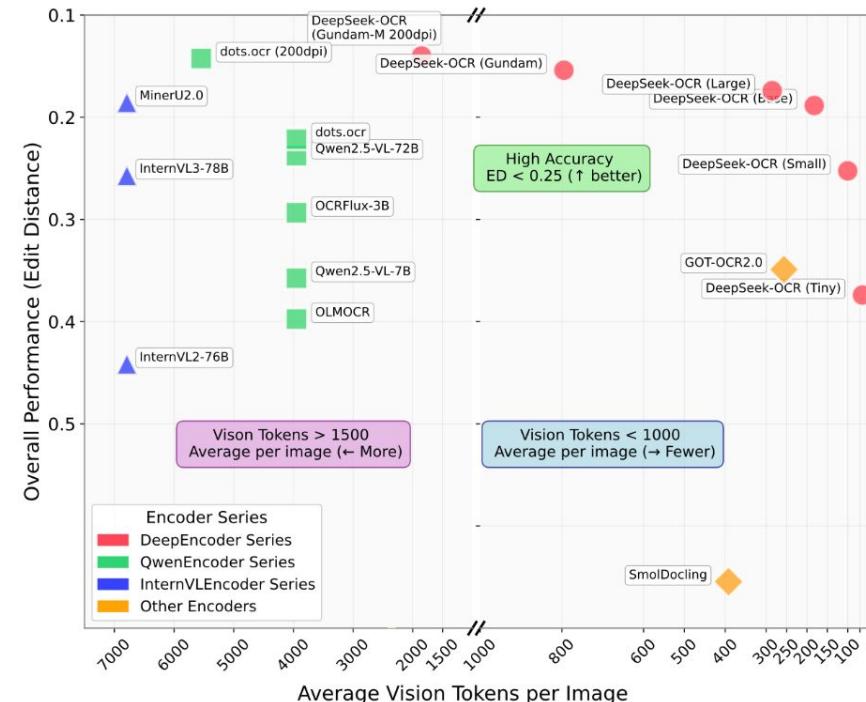
Risks:

- Hallucination
- Hard to guarantee legal accuracy

Must be used with verification layers



(a) Compression on Fox benchmark



(b) Performance on Omnidocbench

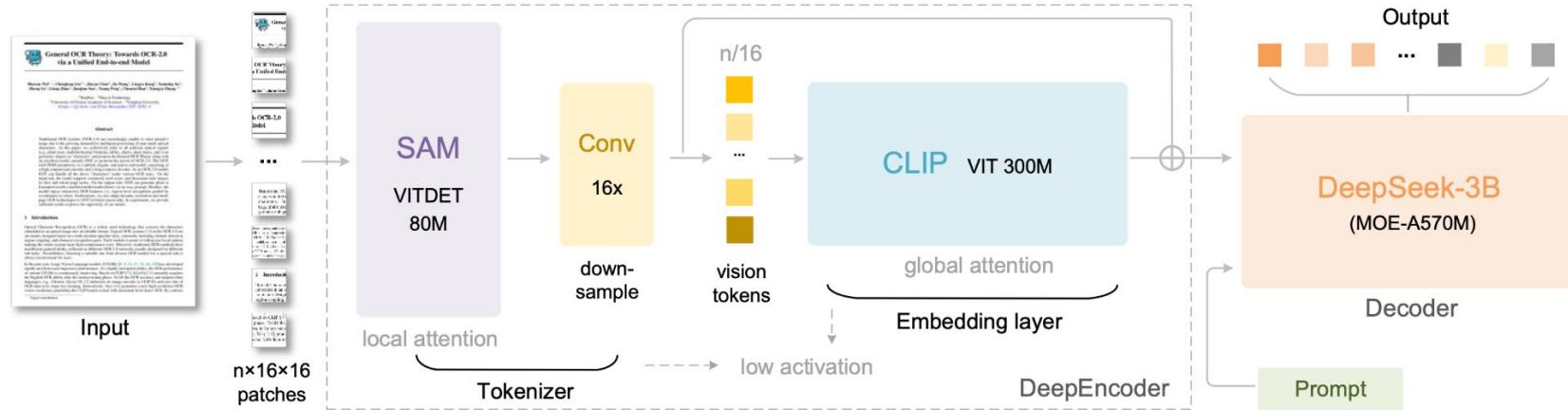


Figure 3 | The architecture of DeepSeek-OCR. DeepSeek-OCR consists of a DeepEncoder and a DeepSeek-3B-MoE decoder. DeepEncoder is the core of DeepSeek-OCR, comprising three components: a SAM [17] for perception dominated by window attention, a CLIP [29] for knowledge with dense global attention, and a $16\times$ token compressor that bridges between them.

<image>\nLocate <|ref|>11-2=<|/ref|> in the image.



<image>\nDescribe this image in detail.



A green plastic container labeled for food storage purposes. The lid is yellow and appears slightly worn out.

On the front lat

- At the top left corner, some Chinese characters.
 - In bold black letters near the center-top reads "6月香" (June Fragrance).
 - Below that, smaller white font states "鮮甜咸香" which translates roughly to "Fresh Sweet Salty Fragrant."
 - Dominating the middle part of the label is large lettering stating "豆瓣醬," meaning "Bean Paste Sauce."

Additional Information includes

- A small circular logo indicating quality assurance ("优质品").
 - An illustration showing what seems like a dish made using bean paste sauce.
 - Text below the main title says "专注25年，只做原酿醋" translating to "Focusing since 1998, only making authentic fermented bean paste."
 - On the bottom right side, additional product specifications include "净含量: 500g".

Overall Style/Ty

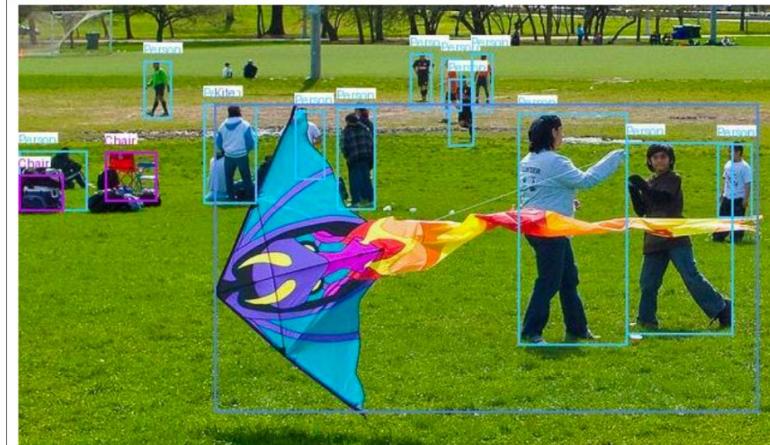
It looks like commercial packaging designed specifically for preserving food items within the jar. It has clear labeling intended for easy identification and understanding among consumers who read Mandarin.

Note: There might be slight variations due to handwriting styles but overall maintains readability regarding brand identity and contents description.

<image>\nLocate <|ref|>the teacher<|/ref|> in the image.



<image>\nIdentify all objects in the image and output them in bounding boxes.



LLM Demo: Thai Text in 2026

- Next-gen understanding of Thai text 
- Fast & accurate recognition
- Context-aware insights from any document
- See AI handle Thai language like never before

LLMs are powerful.

They understand, generate, and decide.

But one question remains:

Can we trust them with
customer data, legal files, land deeds?

This is a **government-level concern**.

Build our own LLM to get:

-  Data privacy
-  Compliance
-  Domain intelligence
-  Full control

No leaks. No doubts.

Text Recognition on Land Title Deeds

Exploring the Challenges and Labeling Accuracy

- **EDA (Exploratory Data Analysis):** Identifying complexity in text recognition
- **Labeling Insights:** Correct vs. Incorrect labels – Key to improving accuracy

Let's dive into the future of land document automation!

Thai LLMs for OCR Enhancement

Thai-focused models:

- WangchanBERT
- Typhoon
- SEA-LION

Use cases:

- OCR post-correction
- Name normalization
- Legal terminology validation

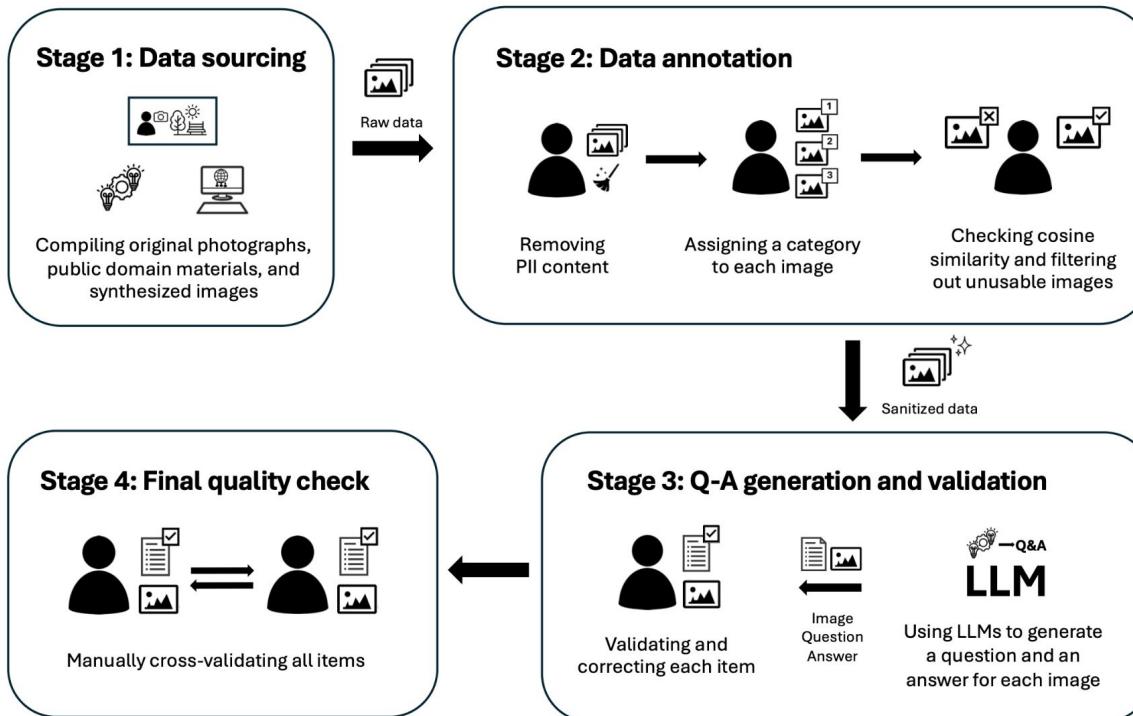
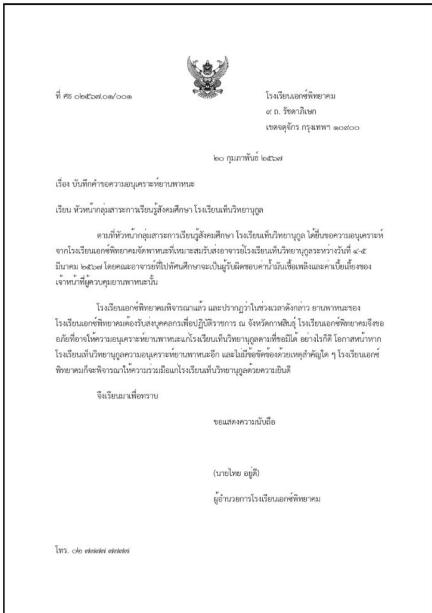


Figure 1: Overview of the ThaiOCRBench data collection and annotation pipeline.

Document parsing



Question

แปลงภาพเอกสารนี้ให้เป็น markdown

Answer

ที่ ศธ ๐๒๕๖๗.๐๑/๐๐๑

<figure>

</figure>

โรงเรียนเอกซ์พิทัยาคม

๙ ณ. รัชดาภิเษก

ເບີນຕຸ້ນຈັກ ກຽມທະພາ ຕອນ

๒๐ กุมภาพันธ์ ๒๐

เรื่อง บันทึกคำขอความอนุเคราะห์ยานพาหนะ

เรียน หัวหน้ากลุ่มส

ตามที่หัวหน้ากลุ่มสาระการเรียนรู้สังคมศึกษา โรงเรียนเทนวิทยานุกูล ได้ขอ

จากโรงเรียนเอกซ์พีทิวคัมจัดพาหนะที่เหมาะสมรับส่งอาจารย์โรงเรียนเท็นวิทยานุกูลระหว่างวัน

มีนาคม ๒๕๖๗ โดยคณะกรรมการที่ไปพิทักษ์ศึกษาจะเป็นผู้รับผิดชอบค่าน้ำมันเชื้อเพลิงและค่าเบี้ยเลี้ยงของเจ้าหน้าที่ผู้ควบคุมยานพาหนะนั้น

โรงเรียนเอกซ์พิทัยาคมพิจารณาแล้ว และปรากฏว่าในช่วงเวลาดังกล่าว ยานพาหนะของ

ໄຊເຮັດອາຄົມທີ່ພະຍານດີວ່າຈະບໍ່ໄດ້ກຳໄຫວ້າໃນລົງທະບຽນ ຈຶ່ງວິທີກຳໄຫວ້າສິນໄລ ໂຊງເຮັດອາຄົມທີ່ພະຍານຈຶ່ງວິທີກຳໄຫວ້າສິນໄລ ອັກນີ້ທີ່ໄດ້ກຳໄຫວ້າຄວາມອຸປະກອນທີ່ບໍ່ມີຢູ່ໃນລົງທະບຽນ ໄດ້ກຳໄຫວ້າສິນໄລ ໂຊງເຮັດອາຄົມທີ່ພະຍານຈຶ່ງວິທີກຳໄຫວ້າສິນໄລ ແລະ ເນັ້ນມີຫຼັກຂໍ້ຕົວຢ່າງທີ່ສັກປູໄລ ໂຊງເຮັດອາຄົມທີ່ພະຍານຈຶ່ງວິທີກຳໄຫວ້າສິນໄລ

WMO/TIROS/Nimbus 1865

กิตติมศักดิ์

(ก่อนปี พ.ศ. ๒๕๖๓)

ผู้อุปถัมภ์ฯ

105 106 107 108 109

Document classification

Question

เอกสารในภาพเป็นประเภทใด ระหว่าง [เอกสารแบบฟอร์มราชการ', 'เอกสารราชการ', 'ทะเบียนสมรส', 'เอกสารการลงบันทึกประจำวันของตำรวจ', 'รายงานทางวิทยาศาสตร์', 'แบบฟอร์มภาษีเงินได้', 'ใบอนุญาตขับรถ', 'ประชณีบัตร']

Answer

แบบฟอร์มภาษาอังกฤษได้

Key information extraction



Question

รายการเครื่องดื่มในเมนูนี้มีอะไรบ้าง ราคาเท่าไหร่ จากรปภกรุงเทพดีง

ข้อมูลมาในรูปแบบ JSON Dictionary ตัวอย่างเช่น:

{"ชื่อ": "ราดา", ...}

Answer

{ "ເອສພຣສ່ (Single)": "55 ບາທ",
"ເອສພຣສ່ (Double)": "65 ບາທ",
"ມອຄາຊື້ນີ້ (ຮ່ອນ)": "75 ບາທ",
"ມອຄາຊື້ນີ້ (ເມື່ອ)": "85 ບາທ",
"ກຳປູໄປ້ນີ້ (ຮ່ອນ)": "75 ບາທ",
"ກຳປູໄປ້ນີ້ (ເມື່ອ)": "85 ບາທ",
"ກາງາມເມລັກຕົວດີ (ຮ່ອນ)": "85 ບາທ",
"ກາງາມເມລັກຕົວດີ (ເມື່ອ)": "95 ບາທ",
"ຈາດຕີ (ຮ່ອນ)": "75 ບາທ",
"ຈາຕີ (ເມື່ອ)": "85 ບາທ",
"ຍອດີ (ຮ່ອນ)": "85 ບາທ",
"ອະເກີກາໄນ (ຮ່ອນ)": "75 ບາທ",
"ບັນວິກາໄນ (ເມື່ອ)": "85 ບາທ",
"ໂລ່ຍ້ຳ (ເມື່ອ)": "85 ບາທ",
"໇າເຫຼືອນົມຄອກຄາວີ້ວັດ": "75 ບາທ",
"໇າເຫຼືອນົມຄອກຄາເທັ້ນ": "85 ບາທ",
"໇າຫຼວມໃຈໆ": "65 ບາທ",
"໇າຫຼັນ (ເມື່ອ)": "85 ບາທ",
"໇າດີເກີນ (ເມື່ອ)": "85 ບາທ",
"ໂລ່ຍ້ຳນອນທີ (ເມື່ອ)": "85 ບາທ",
"ໂລ່ຍ້ຳນັບຄະລິຍົມ (ເມື່ອ)": "85 ບາທ",
"ບັນລຸ ສປັບປຸງ ພັ້ນ (ເມື່ອ)": "115 ບາທ",
"ກາລິນ ໂກສີເລື້ອນ (ເມື່ອ)": "95 ບາທ"

Key information mapping



Question

ตามข้อมูลในภาพ กรุณاجับคุ้มและค่าที่เกี่ยวข้องดังต่อไปนี้: 'ปริมาณพลังงาน', 'คาร์โบไฮเดรต', 'ไขอาหาร', 'วิตามินซี', 'ไขมันรวม' และ 'โซเดียม' กับค่าที่จะถูกจับคู่ได้แก่ '50 kcal', '0 mg.', '2 g.', '40%', '0 g.' และ '135 มก.' เลือกรายละเอียดจากนี้อ่อนไหวในเอกสารแล้วให้คำตอบกลับมาในรูปแบบ JSON โดยตรง:

Answer

{

```
"ปริมาณพลังงาน": ["50 kcal"],  
"ไขมันรวม": ["0 g."],  
"คาร์โบไฮเดรต": ["0 mg."],  
"ไขอาหาร": ["2 g."],  
"วิตามินซี": ["40%"],  
"โซเดียม": ["135 มก."]
```

}

Figure 6: Examples of ThaiOCRBench: (Top) Infographics VQA, (Bottom) Key information mapping

Table parsing



Question

จะแปลงตารางทั้งหมดให้รูปภาพให้เป็นรูปแบบ HTML

Answer

สาขา-เขต(Branch)	ทะเบียนผู้ใช้
น้ำ (Account no.)	เส้นทาง-ลำดับ (Route)
เลขที่แจ้งค่าน้ำ (Invoice no.)	วันที่แจ้งค่าน้ำ (Invoice date)
12-02	
534-400	017690-9
03/01/68	
09:42	วันที่อ่านครั้งนี้ (Current date)
เลขในมาตรา (Current reading)	วันที่อ่าน
ครั้งก่อน (Previous date)	ครั้งก่อน (Previous reading)
จำนวนน้ำใช้ (Consumption)	
03/01/68	2600
2574	03/12/67
26	
ค่าน้ำดิบ (ลูกบาศก์เมตรละ...0.15...บาท)	
3.90	ค่าน้ำประปา R1(00)

Cognition VQA



Question

ราคาของตู้เย็น Beko RDNT252I50HFK ปกติกี่บาท

Answer

["9,490 บาท", "9,490"]

Figure 13: Examples of ThaiOCRBench: (Top) Handwritten content extraction; (Bottom) Cognition VQA

Human-in-the-Loop System

For legal documents:

- AI assists
- Human verifies
- System learns from corrections

AI augments officers, not replaces them

Key Takeaways for Land Department

- OCR is a **system**, not just a model
- Thai language requires special care
- Modern AI improves robustness
- Accuracy and trust are more important than speed

Practical AI = Safe, Auditable, and Scalable

Let's Do the Lab

Part 1: Text Detection

- Test **CHULA** instantly with YOLOv8, v11, v12+ on diverse datasets
- Demos:
 - 👉 Medical Pills Dataset
 - 👉 African Wildlife Detection 



Part 2: Text Recognition

- Try **PaddleOCR**, **EasyOCR**, **Tesseract**
- Explore **LLM-powered OCR**



⚠ Disclaimer:

This notebook is for **educational purposes only** to teach **responsible use of AI for OCR and text recognition**. It is **not intended for commercial or illegal use** and should be used strictly for **academic learning**.

Part 3: Mini-Project

- Apply OCR for **text analytics on documents**

Practical AI for Text Detection and Recognition in Land Title Deed Documents
Teerapong Panboonyuen

Thank you ·
Questions welcome



✨ Slide update on January 6, 2026