

# DAGRI subtask 3: 栽培暦を対象とした Figure and Tables Question Answering (FiT-QA) の提案

中川 董(小樽商科大学) 木村 泰知(小樽商科大学) 會田 勇斗(ストックマーク株式会社)  
高橋 洸丞(ストックマーク株式会社) 門脇 一真(株式会社日本総合研究所)  
小林 暁雄 大友 将宏 石原 潤一 馬場 研太 桂樹 哲雄 (農研機構 農業情報研究センター)

## 背景・目的

農業分野では、作物ごとの作業時期や栽培管理の指針を体系的に示した**栽培暦**が用いられている  
現在は、各自治体がこれらの文書を作成、営農指導員が個別に説明を行っている

栽培暦は、**テキスト・図表・画像が混在**しており、マルチモーダルLLMにおいても十分に正確な読み取り・応答ができない  
また、LLMに学習されている**日本の農業知識も十分でない**

## 目的

- 栽培暦を対象としたマルチモーダル質問応答タスクを定義する
- 農業従事者が栽培暦から必要な情報へより容易に到達できるようにするための技術的基盤の構築を目指す

## 提案タスク

※本タスクは、国立情報学研究所 (NII) が主催する評価型ワークショップ NTCIR-19に採択されたタスク DAGRI のサブタスクの一つ

### 使用するデータセット

人手とLLMを用いて作成された栽培暦のVQAデータセット(詳細はP5-2参照)

### 入力

- 質問文(「幼穂形成期は何月の何日?」)
- 画像(栽培暦のPNGデータ)

### 出力

質問に対する回答(「7月中旬」)

### 評価

BLEUScore, LLM-as-a-Judge

### データフォーマット (Trainデータ/Testデータ)

json形式で作成

- Trainデータでのみ提供するデータ
- Test/Trainデータ共通で提供するデータ

question_id	質問に対して連番で付与される固有の番号
file_name	画像ファイルのファイル名
page	対象画像の該当する文書中のページ番号
question	質問文

image_width	画像ファイルの横幅サイズ
image_height	画像ファイルの縦幅サイズ
bbox	回答の根拠となる箇所のbbox座標
answer	正答
synthetic	LLMによる自動合成QAを採用したか
difficulty	難易度(今回のデータセットは"easy")

栽培暦

## ベースライン手法による評価

### ベースライン手法

質問文と画像を入力として、既存のマルチモーダルLLMからの出力を得る  
BLEUScoreで評価 (Train/Testデータ合計687件を対象)

### 使用モデル

- GPT-5.2 (gpt-5.2-2025-12-11)
- Qwen3-VL-8B-Instruct

### システムプロンプト・設定

Answer the question using only the information shown in the image.  
Follow these rules:  
- The answer must be a short phrase or single word. Do not use full sentences.  
- Use the minimum number of words necessary.  
- If multiple pieces of information are required, list them using the Japanese comma "、".

temperature = 0.7, top\_p = 0.9

### 評価

Model	BLEU ※文字ベース	BLEU ※Mecabベース
gpt-5.2-2025-12-11	57.66	40.90
Qwen3-VL-8B-Instruct	65.17	35.54

正解とモデルの出力それぞれについて

- Unicode 正規化
- "~", "-"の文字コードの正規化
- を行い、fugashiと unidic-liteを用いてトークナイズ

### 誤り分析

時系列に関する質問での誤答が多くを占める

- 正解よりも幅広い期間を答える
- 上旬、中旬などのずれ

他に

画像が複雑であることで、抽出箇所を誤るケースが見られた

### おわりに

- 本研究で提案する FiT-QAは、データ資源の少ない日本の農業の暦に関する質問応答タスクをテキストと画像というマルチモーダルな設定で提案する
- ベースライン手法による評価では、既存のモデルでは完全に解き切ることができないタスクとなっている

Q: 収穫が行われるのは何月何日からいつまでですか?  
A: 7月下旬~8月上旬  
例: 「7月中旬~8月上旬」「7月下旬~8月中旬」