

1 PNAS Template for Supporting Information

2 This PNAS template for Supporting Information (SI) may be used to organize your supporting material. **Once formatted,**
3 **this first page should be deleted by removing the \instructionspage command.** The template is intended to provide
4 a clearly organized PDF file that will ensure readers can easily navigate to sections or specific figures and tables. Movie files or
5 large datasets can be presented as separate files. Further information is available in the [PNAS Author Center](#).

6 Using the template

7 Specify the title, author list, and corresponding authors with the `\title`, `\author` and `\correspondingauthor` commands.
8 The cover page will be automatically generated with the relevant description of the SI, by the `\maketitle` command.

9 Figures should be placed on separate pages with legends set immediately below each figure. Table titles should be set
10 immediately above each table. Note that tables extending beyond the width of the page can be included in the PDF or provided
11 as separate dataset files. Oversized/nonstandard page sizes are accepted as part of your SI Appendix file.

12 References cited in the SI text should be included in a separate reference list at the end of this SI file: (1) and (2).

13 Supporting information for Brief Reports is limited to extended methods, essential supporting datasets, and videos (no
14 additional tables or figures). Supporting figures and tables are not allowed for Brief Reports.

15 Submitting SI

16 Delete this first page by removing the `\instructionspage` command, and then save your completed SI file as a PDF for
17 submission. Further submission instructions are available [here](#).



18

19 **Supporting Information for**

20 **Orchestrator-Agent Trust: An Agentic AI Visual Classification System with Confidence-Aware** 21 **Orchestration and RAG Reasoning**

22 **Konstantinos I. Roumeliotis, Ranjan Sapkota, Manoj Karkee and Nikolaos D. Tselikas**

23 **Corresponding Authors.**

24 **E-mail: k.roumeliotis@uop.gr, rs2672@cornell.edu**

25 **This PDF file includes:**

- 26 Supporting text
- 27 Figs. S1 to S5
- 28 Tables S1 to S2
- 29 Legends for Movies S1 to S3
- 30 Legends for Dataset S1 to S2
- 31 SI References

32 **Other supporting materials for this manuscript include the following:**

- 33 Movies S1 to S3
- 34 Datasets S1 to S2

Table S1. Summary statistics of confidence scores for each agent during zero-shot classification (Experiment 1).

Statistic	Qwen	GPT-4o	Orchestrator
Count	160	160	160
Mean	0.943	0.874	0.917
Std. Dev.	0.042	0.042	0.037
Min	0.800	0.700	0.825
25th Pctl	0.950	0.850	0.900
Median	0.950	0.850	0.925
75th Pctl	0.950	0.900	0.950
Max	1.000	0.950	0.980

Table S2. Comparison of fine-tuning configurations for GPT-4o using ResNet-50–derived and default hyperparameters.

Base Model	Epochs	Batch Size	Val. Loss	Duration (s)	Cost (USD)
GPT-4o (ResNet-50 tuned)	10	16	0.0088	1,778	47.53
GPT-4o (Default settings)	3	1	0.0617	1,652	13.09

35 **Supporting Information Text**

36 **Subhead.** Type or paste text here. This should be additional explanatory text such as an extended technical description of
37 results, full details of mathematical models, etc.

38 **Heading**

39 **Subhead.** Type or paste text here. You may break this section up into subheads as needed (e.g., one section on “Materials”
40 and one on “Methods”).

41 **Materials.** Add a materials subsection if you need to.

42 **Methods.** Add a methods subsection if you need to.

```

"agent_prompt": """"Analyze the provided image of an apple leaf using your
computer vision capabilities. Classify the leaf into the most appropriate
category based on its condition, choosing from the predefined list:
{
  "categories": [
    "black-rot",
    "healthy",
    "rust",
    "scab"
  ]
}
Provide your final classification in the following JSON format without
explanations:
{
  "category": "chosen_category_name",
  "reason": "brief_reason_for_this_classification",
  "confidence": "confidence_score_between_0_and_1"
}""",

"agent_response": {'category': 'scab',
                    'reason': 'The presence of brown spots and scarring on the'
                    'leaf surface is characteristic of scab disease.',
                    'confidence': 0.95, 'time': 4.0999331474, 'cost': 0.001035}

```

Fig. S1. AI Agent Prompt and Response

```

"orchestrator_prompt": """"You are acting as the orchestrator in an agentic AI
system. Your role is to make an informed and well-reasoned final decision
for a classification task, based on the outputs of multiple AI agents.
Task:
Classify an image of an apple plant leaf into the correct disease category.
Context:
You have received independent responses from two specialized AI agents.
Each agent may use different modalities, models, or reasoning approaches.
Your responsibility is to analyze, compare, and synthesize their outputs to
make a final, justified classification decision.

$agents_response

Your Task as Orchestrator:
Based on the agents' responses, provide your final classification decision.

You must return your answer in the following JSON format:
{
  "category": "chosen_category_name",
  "reason": "detailed explanation of your reasoning process",
  "confidence": "confidence_score_between_0_and_1"
}

Be sure your reasoning includes:
- How you evaluated the agents' outputs
- Any conflicts or agreements you observed
- The rationale behind your chosen decision""""

"orchestrator_response": {'category': 'scab', 'reason': 'Both AI Agent 1 and'
  'AI Agent 2 identified the disease as scab based on the presence of dark,'
  'scabby spots and characteristic brown spots with scarring on the leaf'
  'surface. Given the strong agreement between both agents, and their high'
  'confidence levels (0.95 and 0.85), the evidence clearly points to scab'
  'disease. This consensus, along with the corroborating descriptions, supports'
  'a reliable classification.',
  'confidence': 0.93, 'time': 3.1174688339, 'cost': 0.000787}

```

Fig. S2. Orchestrator Prompt and Response

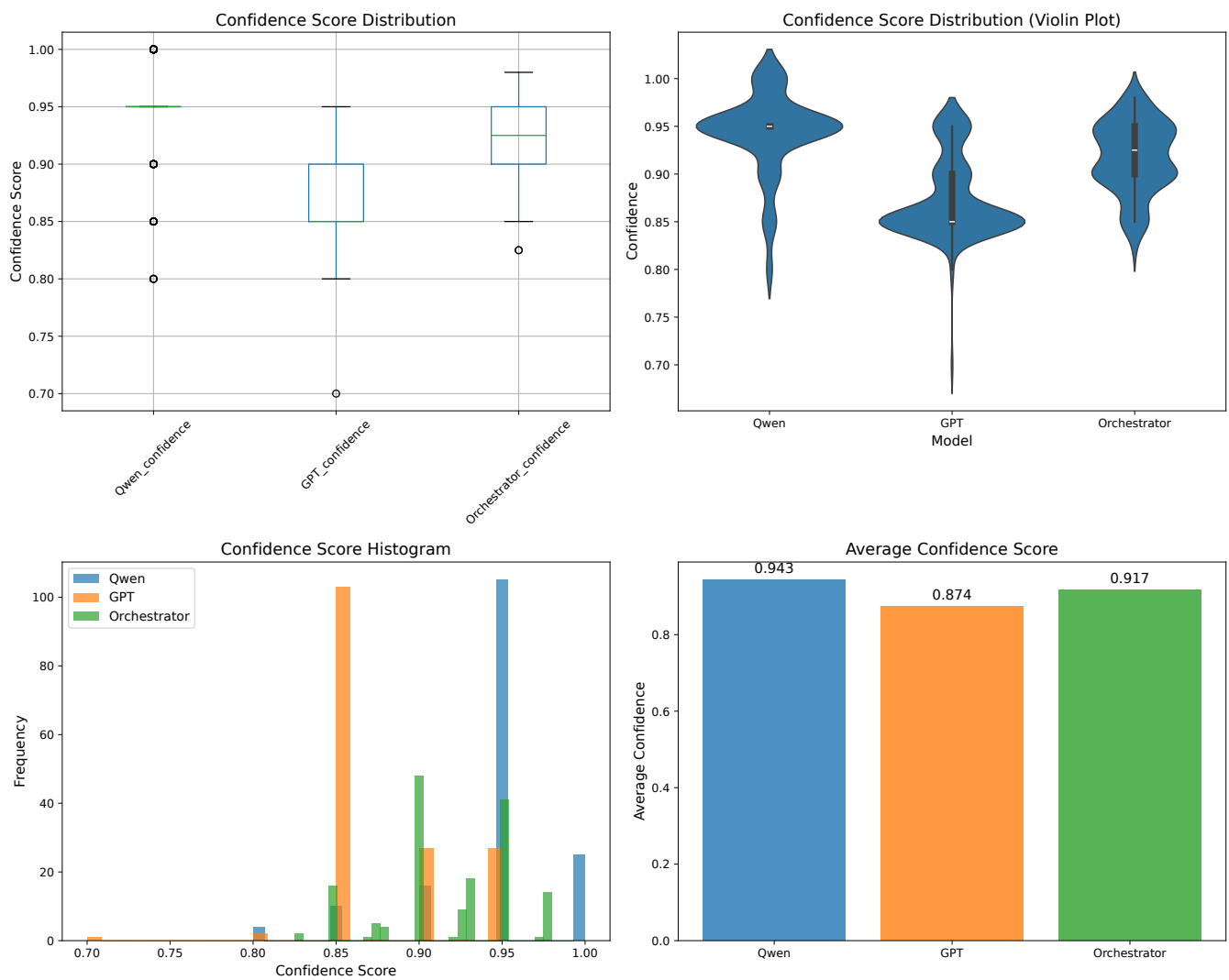


Fig. S3. Experiment 1. Confidence Analysis

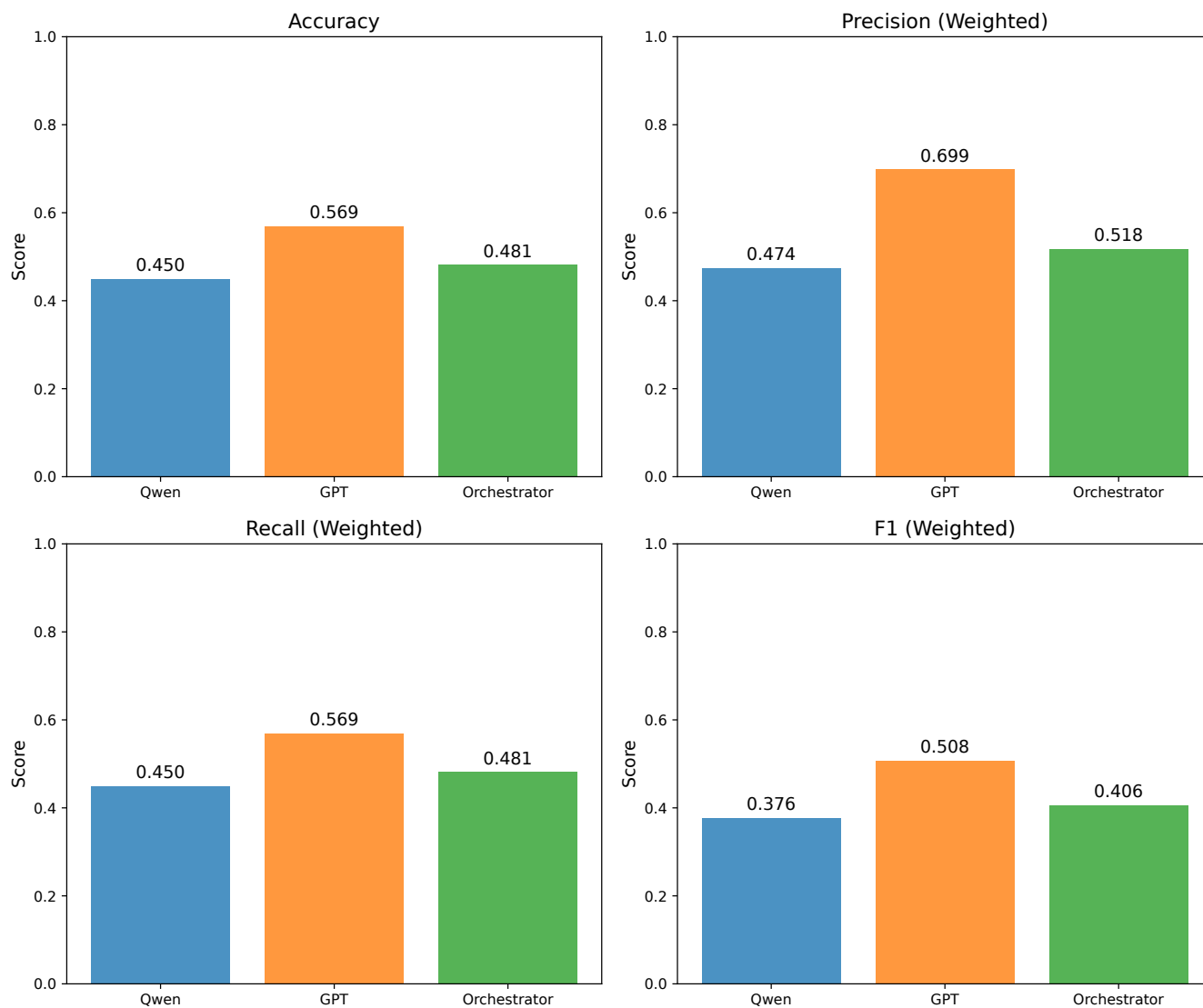


Fig. S4. Experiment 1. Classification Metrics Comparison

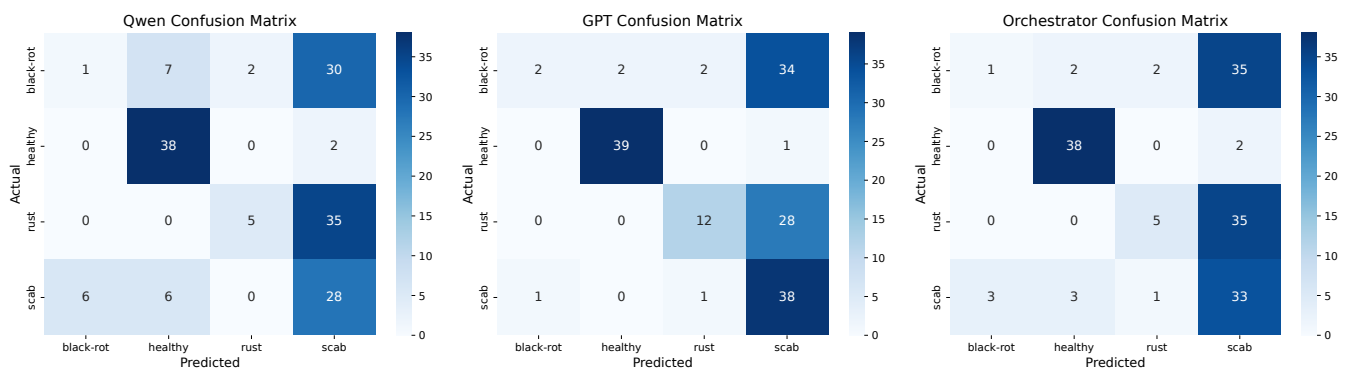


Fig. S5. Experiment 1. Confusion Matrices

43 **Movie S1.** Type legend for the movie here.

44 **Movie S2.** Type legend for the other movie here. Adding longer text to show what happens, to decide on
45 alignment and/or indentations.

46 **Movie S3.** A third movie, just for kicks.

47 **SI Dataset S1 (dataset_one.txt)**

48 Type or paste legend here.

49 **SI Dataset S2 (dataset_two.txt)**

50 Type or paste legend here. Adding longer text to show what happens, to decide on alignment and/or indentations for
51 multi-line or paragraph captions.

52 **References**

- 53 1. A Varga, AN Edmonds, Multilingual extraction and editing of concept strings for the legal domain. *Adv. Comput. Sci. an*
54 *Int. J.* **5**, 18–23 (2016).
- 55 2. TE Olsen, G Stensland, On optimal timing of investment when cost components are additive and follow geometric diffusions.
56 *J. economic dynamics control* **16**, 39–51 (1992).