

TABLE I: Comparison of *LitAI*, Salesforce BLIP, GPT-4, and Google Gemini on figure interpretation for Figures 1 and 2

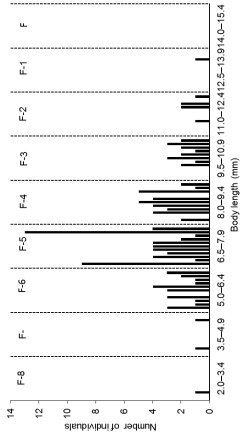
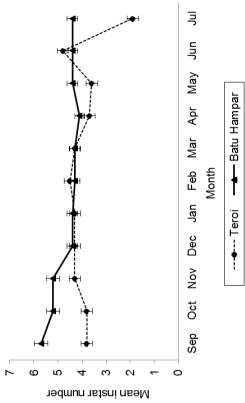
Sample Image	Salesforce Image captioning	GPT 4	Google Gemini	<i>LitAI</i> (ours)
	<p>a plot of the number of the different species in the species</p>	<p>The image is a histogram displaying the distribution of individuals across various ranges of body length measured in millimeters. The x-axis represents the body length categories, ranging from 2.0 mm to 15.4 mm, divided into increments. The y-axis indicates the number of individuals falling within each body length category. Vertical bars represent the count of individuals for each size range, with labels above the x-axis denoting different sample sets or groups (F-1 to F-8 and F). This type of visualization is typically used in biological or ecological studies to analyze size distribution within a population.</p>	<p>The image you sent is a graph showing the distribution of body lengths in a population of animals. The x-axis is labeled "Body length (mm)" and the y-axis is labeled "Number of individuals". The graph shows that there are more animals with body lengths in the middle range (around 7.0-9.4 mm) than there are animals with shorter or longer bodies.</p>	<p>This image presents a bar graph detailing the monthly collection counts of individual <i>Thalerosphyrus</i> nymphs, categorized by body length in millimeters (mm). The x-axis delineates the body length ranges, while the y-axis quantifies the number of individuals. The graph organizes the data into categories labeled F-8 through F, presumably corresponding to different instar stages of <i>Thalerosphyrus</i> nymphs. Each bar within a category illustrates the count of nymphs falling within specific body length ranges. The purpose of the graph is to depict the segregation of <i>Thalerosphyrus</i> nymphs into various instar stages based on body length. Such analysis is instrumental in exploring the growth patterns and developmental progressions of the species over time.</p>
	<p>a line graph of the average and average time of a person</p>	<p>The graph depicts the monthly trend in the mean instar number of two groups or species labeled as "Terai" and "Batu Hampar" over a period from September to July. The y-axis represents the mean instar number, ranging from 1 to 7, while the x-axis displays the months. The "Terai" group is shown with a dashed line and diamond markers, indicating a generally decreasing trend from September, stabilizing around the number 4 from February to June, and then showing a sharp decline in July. On the other hand, "Batu Hampar," represented with a solid line and triangle markers, exhibits a relatively steady pattern, fluctuating slightly around the mean instar number 4 throughout the year. Error bars on both lines suggest variability in the data, although it remains relatively consistent for "Batu Hampar." This graph likely reflects some biological or ecological study observing developmental stages or conditions of these groups across different times of the year.</p>	<p>The image you sent is a line graph showing the average number of visitors per month to two tourist attractions, Teroi and Batu Hampar, over a year. The x-axis of the graph shows the month, from September to the following June. The y-axis shows the average number of visitors. The line for Teroi is blue and the line for Batu Hampar is green. The data for Teroi shows that the highest average number of visitors is in December and January. The data for Batu Hampar shows that the highest average number of visitors is in April, May and June. It is difficult to say from the graph which tourist attraction is more popular overall, but Teroi appears to be more popular in the winter months, while Batu Hampar appears to be more popular in the spring and summer months.</p>	<p>The image is a line graph showing the growth of <i>Thalerosphyrus</i> populations in two rivers, Teroi and Batu Hampar, over the course of several months. The Y-axis represents the mean instar number, with higher numbers indicating younger instars, and the X-axis represents the months from September to July. The graph shows that the <i>Thalerosphyrus</i> population grew at a similar rate in both rivers, with some fluctuations throughout the months. The mean instar number generally decreases over time, indicating the aging of the population as the instars mature.</p>

TABLE II: Comparison of *LitAI*, Salesforce BLIP, GPT-4, and Google Gemini on figure interpretation for Figure 3

Sample Image	Salesforce Image captioning	GPT 4	Google Gemini	<i>LitAI</i> (ours)
	<p>a diagram of the process of the human body's organ</p>	<p>The diagram provides an integrated view of the impact of COVID-19 on the body, particularly focusing on the cytokine storm's effect on heart health. It starts with COVID-19 infection in the lungs and the subsequent cytokine release, which includes IL-6, IL-8, and MCP-1, leading to systemic effects such as fibrosis, thrombosis, and reduced heart contractility. The lower section outlines the experimental approach using organoid models, specifically cardiac organoids stimulated with IL-1β to mimic and study the cytokine storm.</p>	<p>This scientific diagram investigates how the SARS-CoV-2 virus impacts the heart using organoids, 3D structures mimicking heart tissue grown from stem cells. Researchers create these organoids, expose them to immune system signaling molecules linked to cytokine storm, and then analyze their shape, gene activity, and protein levels to understand how COVID-19 infection might cause heart damage.</p>	<p>This image illustrates a schematic of research into the effects of COVID-19 on the human body, with a focus on the release of cytokines and their impact on the heart. The diagram showcases several components:</p> <ol style="list-style-type: none"> 1) At the top left, the COVID-19 virus is shown infecting the lungs. 2) The release of cytokines is then depicted with blue dots flowing from the lungs to the heart. The diagram annotates three types of cytokines: IL-6, IL-8, and MCP-1, and links their overproduction to a cytokine storm. 3) The consequences of a cytokine storm are presented as fibrosis, thrombosis, and decreased contractility in the heart. <p>Below this part, the image outlines a research methodology titled "Organoid Modeling of Cytokine Storm and Data Analysis". This section includes several steps:</p> <ol style="list-style-type: none"> 1) Organoid Fabrication: It starts with the creation of an organoid in a petri dish. 2) IL-1β Stimulation: This organoid is then stimulated with IL-1β (a pro-inflammatory cytokine) to model the cytokine storm. 3) Cardiac Organoid: The resultant cardiac organoid is shown, representing a heart model for studying the effects of cytokines. 4) A set of equipment and analysis methods are detailed: a Cytokine Plex device for measurement of cytokines, immunofluorescence for morphological analysis, RNA sequencing, and a fractional area change. <p>This schematic serves to provide a detailed and visual representation of ongoing research into the cardiovascular implications of COVID-19, utilizing advanced organoid models.</p>

TABLE III: Comparison of *LitAI*, Salesforce BLIP, GPT-4, and Google Gemini on figure interpretation for Figure 4

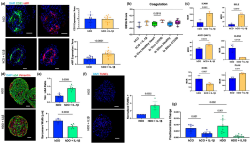
Sample Image	Salesforce Image captioning	GPT 4	Google Gemini	<i>LitAI</i> (ours)
	<p>a diagram of the different cells in the human cell are shown</p>	<p>The image presents various experimental data and microscopy images related to cellular responses under different conditions, most likely in the context of heart tissue research. Panel (a) shows fluorescence images comparing the expression of CD31 and vWF under control conditions and with IL-1β treatment, alongside quantitative bar graphs. Panel (b) depicts a scatter plot assessing coagulation scores across different conditions, including healthy and COVID scenarios. Panel (c) displays bar graphs comparing mRNA expression levels of various genes (ICAM1, SELE, AOC3, CLDN5, NOS3, PLAU) under control and IL-1β stimulated conditions. Panel (d) shows histological staining for αSA and Vimentin in heart organoids with and without IL-1β. Panel (e) and (f) present quantitative analyses of αSA to Vimentin ratio and TUNEL assay results respectively, indicating changes in sarcomere width and apoptosis with treatments. Panel (g) shows a scatter plot depicting fractional area changes under different experimental conditions. All panels include significant p-values, suggesting alterations in gene expression, protein levels, and cellular morphology upon IL-1β stimulation or in disease models.</p>	<p>Researchers in this study utilized human heart organoids, grown from stem cells, to investigate how COVID-19 infection damages the heart. They mimicked the cytokine storm, an overreaction of the immune system, by exposing the organoids to specific molecules. By analyzing changes in the organoids' shape, gene activity, and protein levels, they aimed to understand how COVID-19 leads to heart injuries.</p>	<p>This image contains several panels (a-g) of microscopy images and graphs showing the effects of IL-1β treatment on human cardiac organoids (hCOs) and human-derived organoids (hDOs). Panel a shows microscopy images of hCOs with and without IL-1β treatment, stained for DAPI (nuclear marker), CD31 (endothelial cell marker), and vWF (von Willebrand factor, another endothelial cell marker). The graphs below the images indicate the expression levels of CD31 and vWF in the organoids. Panel b shows a graph comparing the coagulation score of hCOs with and without IL-1β treatment, <i>in vivo</i> healthy heart, <i>in vivo</i> COVID, <i>in silico</i> healthy, and <i>in silico</i> COVID. Panel c shows graphs comparing the expression levels of various endothelial cell markers (ICAM1, SELE, AOC3/VAP1, CLDN5, NOS3, and PLAU) in hCOs with and without IL-1β treatment. Panel d shows microscopy images of hDOs with and without IL-1β treatment, stained for DAPI, αSA (sarcomeric alpha-actinin, a cardiac muscle marker), and Vimentin (a mesenchymal marker). Panel e shows graphs comparing the αSA to Vimentin ratio and sarcomere width in hDOs with and without IL-1β treatment.</p>

TABLE IV: Comparison of *LitAI*, Salesforce BLIP, GPT-4, and Google Gemini on figure interpretation for Figures 5 and 6

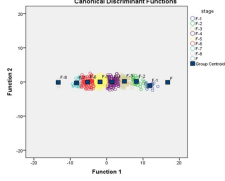
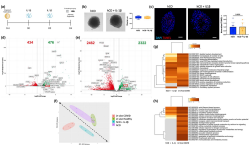
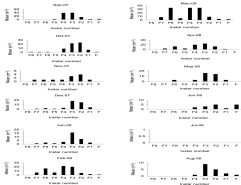
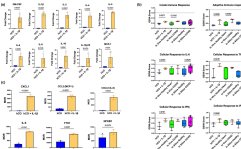
Sample Image	Salesforce Image captioning	GPT 4	Google Gemini	<i>LitAI</i> (ours)
	<p>a plot of a plot of a number of different functions</p>	<p>This image displays a scatter plot titled "Canonical Discriminant Functions," which visually represents a multivariate statistical analysis. The plot shows data points grouped by various stages (F-1 through F-8), each marked by different colors and symbols. The axes, labeled "Function 1" and "Function 2," likely represent the scores of the first two canonical discriminant functions derived from the data, used to maximize the separation among the predefined groups. Group centroids are indicated by squares filled with corresponding group colors, highlighting the average position of each group in this two-dimensional function space. This type of visualization is commonly used in statistics to understand how well different groups can be distinguished from each other based on the functions derived from their characteristics.</p>	<p>The image you sent is a scatter plot showing multiple colored squares. The text labels on the axes indicate that this is a canonical discriminant functions graph. Canonical discriminant analysis is a statistical technique used to mathematically differentiate between two or more groups. The data points plotted as squares represent observations in the dataset, and the axes represent the functions that maximally differentiate between the groups. In this case, the colors of the squares likely correspond to different groups in the data set.</p>	<p>The image shows a canonical discriminant function plot, which is used to visualize the separation of different instar stages (developmental stages) of <i>Thalerosphyrus</i>, an aquatic insect. Each colored circle represents a different instar stage, ranging from F-1 to F-8, and the squares labeled Group Centroid represent the average value for each stage. The plot is divided into two axes, Function 1 and Function 2, which are used to measure the differences between the stages based on body length. The plot shows a clear separation of the different instar stages, indicating that body length is a significant factor in distinguishing between them. This plot is part of a study that analyzed the body length of <i>Thalerosphyrus</i> nymphs collected from two rivers, Batu Hampar and Teroi, and the results are discussed in the corresponding research paper.</p>
	<p>a diagram of the different types of cell lines and their corresponding functions</p>	<p>This image presents a comprehensive set of experimental data and analysis from a biological study involving the effect of interleukin-1 beta (IL-1β) on human colon organoids (hCO). Panel (a) shows a timeline of organoid fabrication and treatment with IL-1β. Panel (b) features images of organoids with and without IL-1β treatment, accompanied by a bar graph comparing their sizes. Panel (c) displays fluorescent images (DAPI, TUNEL) indicating cell death in these organoids, with a corresponding bar graph showing quantification. Panels (d) and (e) are volcano plots illustrating gene expression changes induced by IL-1β, highlighting significant upregulated and downregulated genes. Panel (f) provides a principal component analysis (PCA) plot comparing gene expression profiles of <i>in vivo</i> COVID, <i>in vivo</i> healthy, hCO, and hCO+IL-1β samples, illustrating the similarities and differences among these conditions. Finally, panels (g) and (h) show heatmaps with hierarchical clustering of gene ontologies associated with different biological processes impacted by the treatment, comparing <i>in vitro</i> results with <i>in vivo</i> COVID conditions. This collection of analyses helps understand the cellular responses to inflammation and potential implications for diseases like COVID-19.</p>	<p>The image contains multiple panels showing data related to an experiment on human cardiac organoids and cytokine storms. The top row (a-c) shows the number of cells in each group under various conditions. The middle row (d-e) shows fluorescence microscopy images, likely of the organoids. The text labels for (d) and (e) mention DAPI and TUNEL, which are staining techniques used to identify cell nuclei and cell death, respectively. The bottom row (f-h) shows graphs with gene ontology (GO) terms. Gene ontology is a system for classifying genes based on their function. The graphs show how the treatments affected the expression of genes related to various biological processes. For example, one graph shows changes in genes related to actin filament-based processes and circulatory system processes. Overall, the image suggests that the experiment investigated how cytokine storms affect human cardiac organoids, and how this might be relevant to COVID-19.</p>	<p>The image provides various graphs and data visualizations pertinent to a study on the effects of IL-1β treatment on human cardiac organoids (hCOs) and its potential link to pathological cardiac fibrosis. Figure 3c features an image of blue-stained cardiac organoids with and without IL-1β treatment, alongside a graph showing the quantification of TUNEL-positive cells, indicating the level of cell death in the organoids. Although Figure 4b is not shown, it likely depicts the Gene Set Variation Analysis (GSVA) results, highlighting an increase in GSVA scores for "Fibroblast Activation" and "Collagen Metabolic Process" in treated hCOs and in silico COVID hCOs, as well as COVID-19 heart samples, suggesting induction of pathological cardiac fibrosis by IL-1β treatment. Other panels include volcano plots showing differentially expressed genes, a principal component analysis (PCA) plot comparing <i>in vivo</i> COVID, <i>in vivo</i> healthy, hCO + IL-1β, and hCO samples, and heatmaps with hierarchical clustering showing the enrichment of gene ontology (GO) terms and pathways across these groups, underscoring the clinical relevance of these findings in understanding the molecular mechanisms by which IL-1β treatment may contribute to cardiac fibrosis.</p>

TABLE V: Comparison of *LitAI*, Salesforce BLIP, GPT-4, and Google Gemini on figure interpretation for Figures 7 and 8

Sample Image	Salesforce Image captioning	GPT 4	Google Gemini	<i>LitAI</i> (ours)
	<p>a diagram of the different types of the different numbers of the numbers</p>	<p>The image displays a series of bar graphs, each representing the size-frequency distribution of Thalerosphyrus individuals in the Batu Hampar River, collected monthly from September 2007 to August 2008. Each graph is labeled with a specific month and shows the mean number of individuals per square meter for each instar stage, ranging from F1 to F8. The x-axis of each graph lists the instar numbers from F1 to F8, while the y-axis measures the mean density in individuals per square meter (m^2). The graphs highlight fluctuations in the population across different stages over the year, with higher frequencies observed in the lower instar stages (F1 and F2), and a noticeable decline in the higher instar stages (F3 to F5), indicating possible survival challenges as the Thalerosphyrus progress through their life cycle.</p>	<p>The image you sent appears to be a scientific chart. The chart contains multiple lines and bars across several dates. The dates range from September 2007 to August 2008. The leftmost axis labeled "Instar number" seems to indicate a unit of measurement. The rightmost axis shows values from 0 to 40. It's difficult to say for sure what the chart depicts without more context, but it likely shows a measurement (Instar number) taken at various points in time.</p>	<p>The image shows a series of bar graphs representing the size-frequency distribution of mean Thalerosphyrus in the Batu Hampar River, collected monthly between September 2007 and August 2008. Each bar graph corresponds to a different month and shows the number of individuals in each instar, ranging from instar 1 (F1) to instar 5 (F5), and further stages up to F8, per square meter (m^2). The graphs show a high number of individuals in instar 1 and 2, with a decline in instar 3 and 4, and the largest individuals in instar 5. The data suggests a high mortality rate for Thalerosphyrus during instar 3 and 4, with successful maturation occurring in instar 5.</p>
	<p>a diagram of the different types of cell phone numbers</p>	<p>The image displays a series of bar graphs and box plots representing the results of a study on the immune response in human colon organoids (hCO) with and without interleukin-1 beta (IL-1β) treatment. Panel (a) shows bar graphs of cytokine levels such as GM-CSF, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, and MCP-1, comparing their fold changes between treated and untreated hCOs. The p-values are indicated above each graph, suggesting statistical significance. Panel (b) shows box plots categorized into innate and adaptive immune responses, cellular responses to various cytokines (IL-6, TNFα, IFNγ, and IFNα), each plotted for different experimental settings: hCO, hCO with IL-1β, in vivo healthy, in vivo COVID, and in vitro slice COVID conditions. These plots illustrate variability and comparative response profiles across different settings.</p>	<p>The image contains several graphs and charts, with text labels referring to different medications and immune system responses. For instance, there are labels mentioning IL-4, MCP-1, and IL-12p70, which are all interleukins, a type of cytokine signaling molecule involved in the immune system. The graphs show changes in the levels of these molecules under various conditions. Overall, the image suggests that the experiment investigated how different medications affect the immune system, potentially in the context of heart health.</p>	<p>The image contains a series of graphs and charts depicting the results of a scientific study on the effects of IL-1β treatment on human cardiac organoids (hCOs). The graphs show the fold change in expression of various cytokines and immune response-related genes in hCOs with and without IL-1β treatment. In panel (a), there are bar graphs showing the fold change in expression of cytokines such as GM-CSF, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, and MCP-1. The expression levels are represented by orange bars for hCOs + IL-1β and blue bars for hCOs without IL-1β treatment. In panel (b), there are box plots depicting the GSEA (Gene Set Variation Analysis) scores for different immune responses, including innate and adaptive immune responses, as well as cellular responses to IL-6, TNFα, IFNγ, and IFNα. The scores are compared across different groups, including hCOs + IL-1β, hCOs, <i>in vivo</i> healthy heart, <i>in vitro</i> healthy COVID, <i>in silico</i> COVID, and <i>in silico</i> healthy. In panel (c), there are bar graphs showing the normalized expression levels (in transcripts per million) of genes such as CXCL1, CCL2 (MCP-1), CXCL8 (IL-8), along with several other markers. These graphs illustrate the significant modulation of gene expression induced by IL-1β treatment, providing insights into the molecular mechanisms of cardiac response in human organoids.</p>