# PPT2Poster: Leveraging LLMs and Diffusion Models for Automated Poster Generation

Swapnil Bag (ES22BTECH11034)

Advisor: Prof. Vineeth N Balasubramanian

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

## Abstract


- Converts PowerPoint presentations into concise, well-structured posters with minimal manual effort.
- Applicable in research (e.g., summarizing conference presentations) and professional settings (e.g., business poster creation).
- Eliminates the need for manual design unlike tools like Canva or PowerPoint.
- Addresses a largely unexplored research problem with high practical value.


## Proposed Model

**Content Summarisation** : The PPT is parsed and the infographics like text images, graphs, tables, etc are extracted. This is fed to an LLM to summarise the content and present it in a predefined format consisting of (elements)

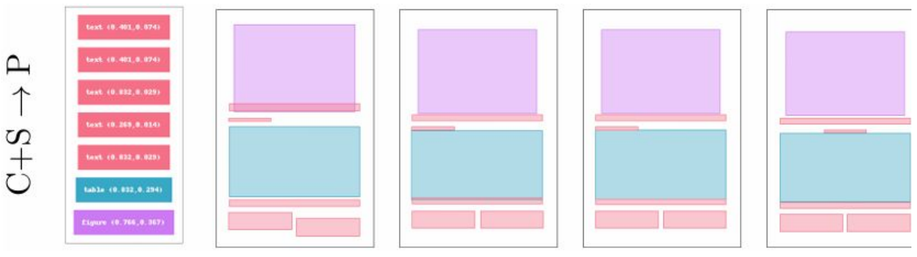
**Layout Generation:** Use Diffusion models (D3PM) techniques to predict a layout conditioned by the elements generated by the LLM. OR Prompt the LLM to generate Bounding boxes

### Input To Model:

$$x = \{(c_1, x_1, y_1, h_1, w_1), \ldots, (c_l, x_l, y_l, h_l, w_l)\}$$

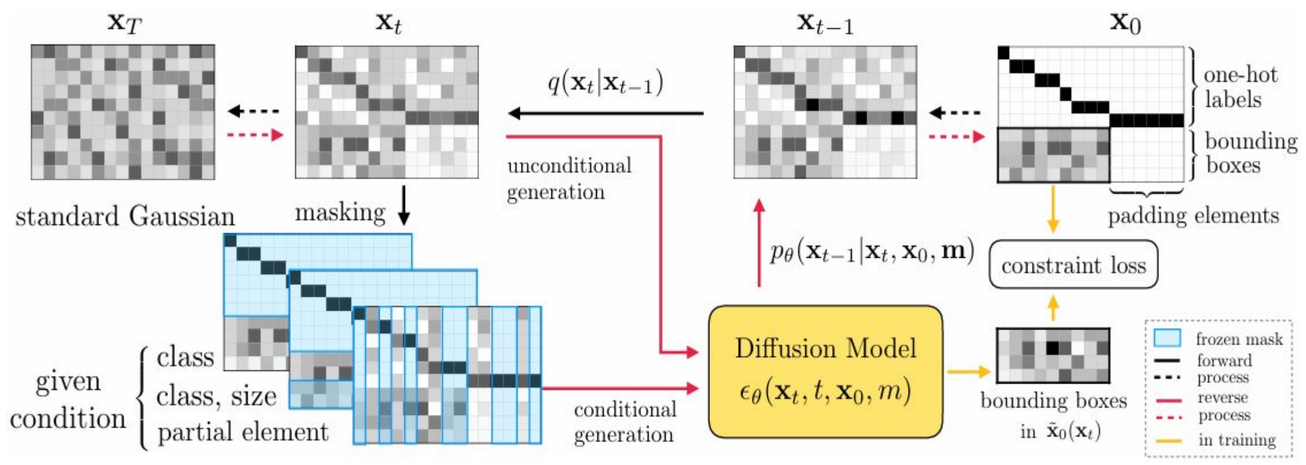
### Conditional mask:

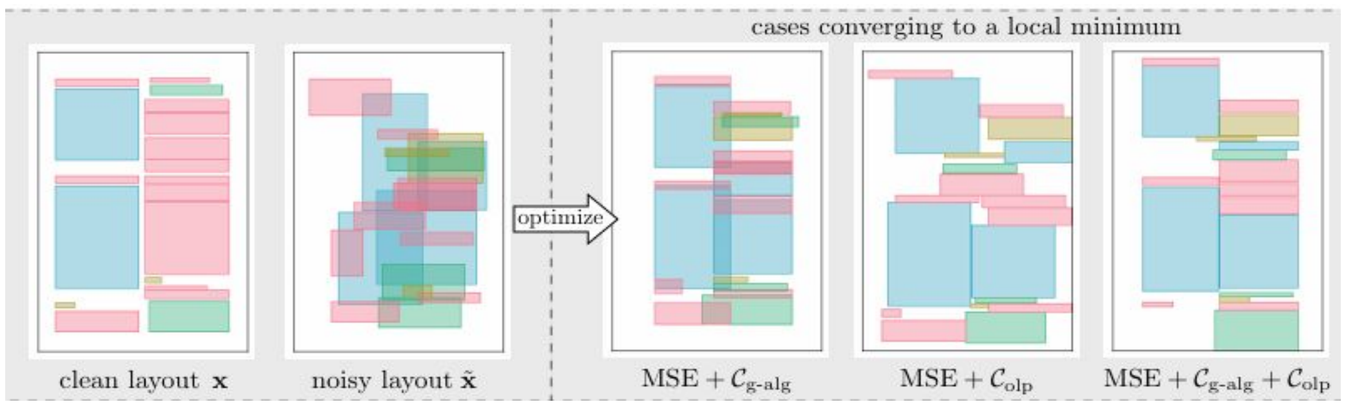
$$\hat{x}_t = m \circ x_0 + (1 - m) \circ x$$

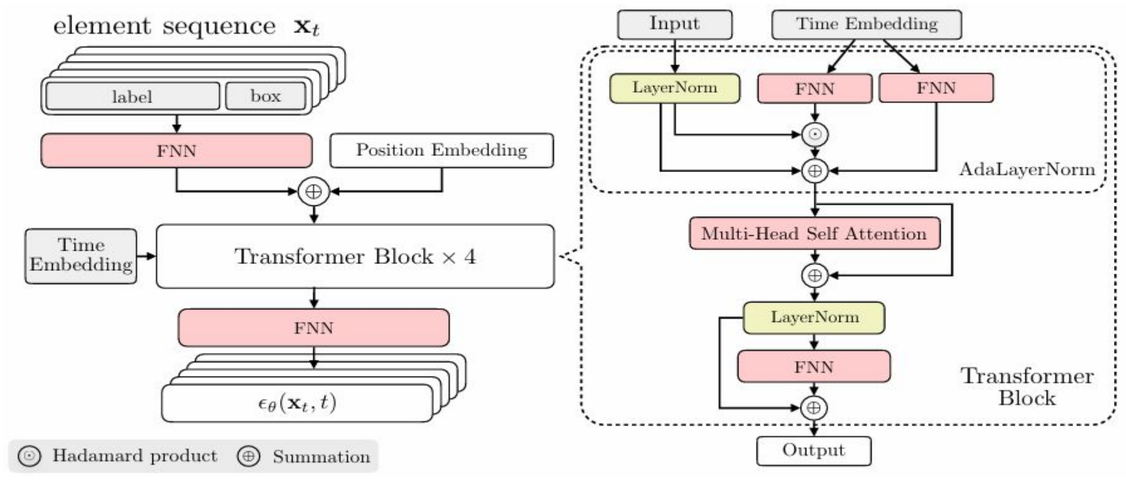



## Bounding Box Generation

### Layout Constraint Diffusion Model (LACE)




## Bounding Boxes using LLM Prompting

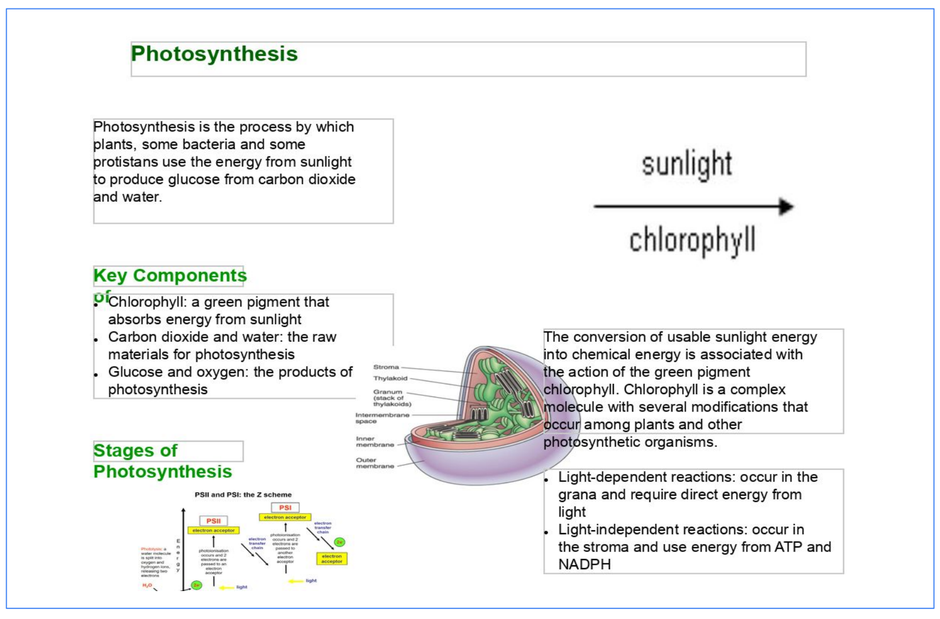
**Algorithm 1** Automated Poster Generation from PowerPoint

1: **Input:** PowerPoint file (slides with text and images)
2: **Output:** Poster layout with arranged content and images
3: Extract all infographic elements from the PowerPoint
4: Summarise the text and decide element types with the content to be displayed (heading, subheading, bullet points, text box, image, etc)
5: Store extracted content in a JSON format with structure: { `"element-type":..., "content":...`}
6: **for** each extracted image **do**
7: Generate a 4–5 line summary describing the image
8: **end for**
9: Pass content and image summaries to the LLM
10: **Prompt 1:** Ask LLM to select the most appropriate image given the selected content
11: **Prompt 2:** Ask LLM to estimate the bounding box $(x, y, h, w)$ for each content block based on size and type
12: Pass content, selected images, and predicted bounding boxes to the poster layout generator
13: Generate final poster with arranged elements

```
Instructions:
- Predict bounding box as [X_center, Y_center, height, width]
- Ensure boxes are:
    - Non-overlapping
    - Aesthetically pleasing
    - Enclose the content
- Heading must be at the top.
- Text boxes and bullet points can be side-by-side or stacked.

Output:
- Strict JSON format with element types as keys.
- No extra text or explanation.

Example:
{
  'heading_1': [600, 70, 50, 900],
  'text_box_1': [450, 300, 300, 400],
  'image_1': [350, 450, 100, 150],
  'bullet_points_1': [250, 400, 400, 300]
}
```




## Future Work

- Currently, we rely on LLM-based prompting strategies for generating bounding boxes. However, due to limitations in spatial reasoning and consistency, the results are suboptimal.
- We plan to explore diffusion-based layout generation methods to achieve more accurate, coherent, and visually appealing poster designs.

## References


- Chen et al. (2023). LACE: Towards aligned layout generation via diffusion model with aesthetic constraints.
- Lin et al. (2023). LayoutPrompter: Awaken the design ability of large language models.
- Zhang et al. (2023). LayoutDiffusion: Improving graphic layout generation by discrete diffusion probabilistic models