Job Description:

To summarise the project:

We're trying to build a system that can take a prompt (and simple layout constraints like a dieline/text zones) and generate a design mockup (specifically focused on packaging as of now) that's good enough to drop into a PDF/EPS file with basic vector layers (e.g., dielines/typography/line art) on top of photo-realistic imagery. Think of it as essentially prompt → image consistent with layout → convert to a hybrid raster+vector asset → assemble into a usable mockup.

This is naturally pretty hard. Getting diffusion/GAN models to respect layout constraints, produce legible text, and then converting the raster output into reasonable vector layers is not trivial. We do not expect a magic "one-click" solution quickly. The value here is in steady progress & finding out which architectures/conditioning tricks work, what to measure, how to convert contours to Bezier paths without garbage artifacts, and where simple, well-placed heuristics beat over-engineering. We've made some progress already (concept note, early prototypes/experiments, a few algorithm papers on conversion and a clear plan for the three work streams below). If you like reading papers, brainstorming, and running small tests to learn what actually works, you'll fit right in and please be open to learning because most probably, most of this will be very new to you (and it was for us too a few months back so don't worry!)

Note: We understand many strong candidates will have no prior experience in this niche; that's fine. It is learning-heavy with substantial reading and brainstorming. Enthusiasm and willingness to learn matter most. We do expect familiarity with basic ML concepts and generative-modeling fundamentals but please don't feel overwhelmed if you are unfamiliar with several terms used in the description below

Role & Basics

- **Role:** Research Intern — Generative Modeling for Design

- **Team Size:** 6 researchers (independent under industry colloboration)

- **Time:** ~6–12 hrs/week; 1 weekly meeting + 1–2 additional syncs as needed

- **Compensation:** None; non-monetary benefits (significant learning, authorship consideration, recommendation letters & industry exposure)

Project Goal

Build an end-to-end pipeline that:

- chooses and trains a core image generator (diffusion or GAN) that can follow layout assets (dielines/text zones) and prompts;

- converts the generated raster into a practical hybrid raster–vector representation (contours → Bezier paths & vector/raster compositing)

- assembles the result into a clean PDF/EPS mockup suitable for downstream design iteration.

Work Streams (pick a primary; collaborate across others)

1) Model Architecture & Brainstorming (Imagination)

- Survey diffusion vs. GAN backbones (latent vs. pixel space) and map trade-offs to our constraints.

- Specify conditioning: prompts + masks/layout assets (dielines/text zones); consider simple ControlNet-style signals.

- Propose objective choices (pixel/feature/edge) and minimal sanity checks that genuinely predict visual quality.

2) Experimentation & Benchmarking (Testing)

- Implement lightweight prototypes and test specific compute blocks (UNet/attention variants, guidance/scaling strategies).

- Run focused ablations on losses and conditioning; evaluate quality vs. speed and layout adherence.

- Keep tight logs/plots; short write-ups of what worked, what failed, and the likely reason.

3) RGB-Raster → CMYK-Vector / Hybrid Conversion

- Segment outputs; extract contours; fit Bézier paths for dielines/typography/line art (aim: clean, few self-intersections).

- Compose vector layers with raster imagery; assemble PDF/EPS with correct ordering and basic metadata.

- Try practical text handling (e.g., OCR → vector outlines) and simple stroke/fill inference where it helps (and it might not, we aren't sure)

Responsibilities

- Read targeted papers and present option sets with clear pros/cons.

- Build, test, and benchmark small code paths; write clean, reviewable code

- Track metrics (quality, adherence to masks/layout, runtime) and visualize failure modes.

- Collaborate with designers/researchers for qualitative checks; document protocols and results.

Minimum Qualifications

- Solid foundations in CS/ML/AI/Math.

- Familiar with at least one of: image-generation algorithms, conversion algorithms, or compression methods.

- Proficiency in Python and fundamental data processing techniques is desirable. While extensive coding may not be a primary initial responsibility, algorithm design, often expressed in pseudocode, will be a regular task. Therefore, specific language requirements are flexible, though familiarity with Python and C++ would be advantageous.

Preferred Qualifications

- Experience with PyTorch (or TensorFlow), Hugging Face Diffusers, and/or OpenCV

- Exposure to vector graphics and basic understanding of how file conversion works (SVG/PDF/JPEG, Bezier curves).

- Prior work with diffusion/GANs, image generative models and basic ML theory **(Highly preferred)**

How to Apply

Fill the ≤5-minute form: https://forms.gle/zLKyusPe2u6HfyCg7 . Questions or to request additional detail: **reply to the outreach email or reach out at +91 7734803500**