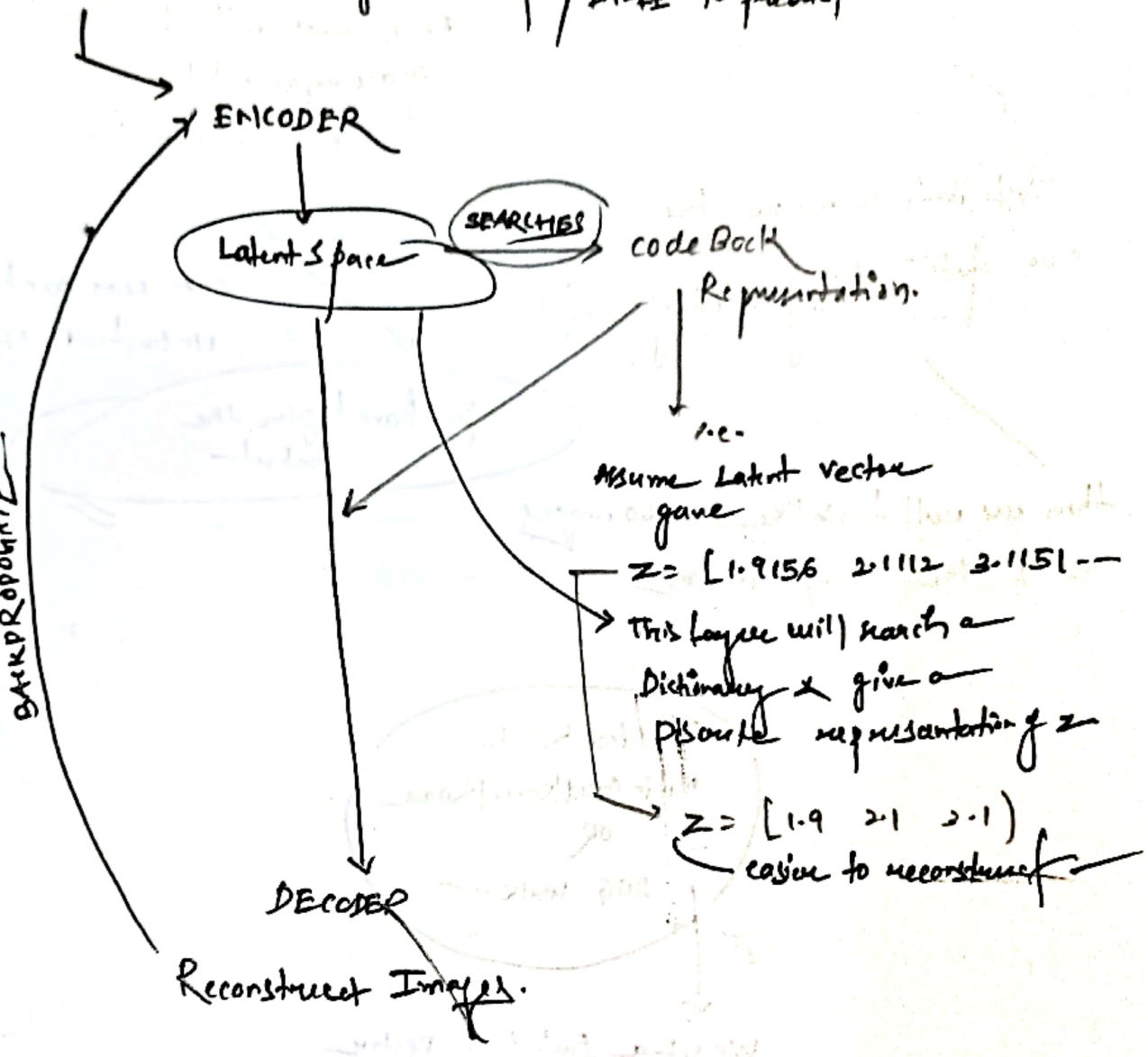
AutoEncoder Architecture

VQ-VAE → used by CHATGPT / DALLE to predict



SYSTEM PROMPT: System prompt Something Like me can that me get 5 commands → Left OR RIGHT, STOP or SPEED up =.

ZERO SHOT

Semantic analyzer Part ??

HOW ??

First initially we will train over AutoEncoder very much efficient

then we will train like 50-60 images with the Unexpected Cases

WHY:

Feed LLM with the Latent Space representation not just the Pixels.

WHY ??

COZ LLM can't just understand this stuff

you have to give the context

Like?

Flooded Road
Nitin Gadkari Road
OR
DOG Walking

We get a Embedded vector with Tags

Now we compare OR COSINE SIMILARITY then it gave Tags to the Image

RAW IMAGE → ENCODED VECTOR → COMPARE

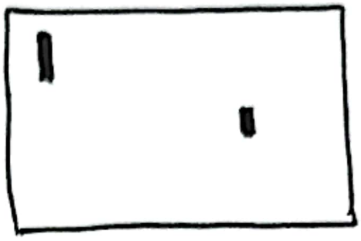
CONTEXT WITH LATENT

LLM

WORKS ??

System Prompt

OR what after getting the content ?!



You know what we can create the occupancy grid representation ?!

↓

Like after is what is the occupancy grid OR is it affecting the path.

FINAL SOLUTION:

GET THE LIDAR DATA REPRESENTATION

HOW?! → DON'T KNOW.