

Button 1 pressed : Text-To-Voice conversion initiates

pressing button 4
Initiates the camera
to capture input

Extracting the text
using Tesseract OCR

Converting Extracted
Text To Voice
GTTS engine

Button 2
pressed : Face-
recognition
initiates

pressing button 4
Initiates the camera
to capture input

Captured Image Send
to cloud for face
recognition task using
FaceNet

Response from Cloud
received to Nav-Voice
which converts it to
Voice using GTTS

Button 3 pressed :
Image Captioning
initiates

pressing button 4
Initiates the camera
to capture input

Captured Image Send
to BLIP image
captioning model on
Hugging face

Response from BLIP
model received and
get converted into
Voice using GTTS

Button 5 pressed :
To send Emergency
alert

Finding real time
device location using
Google cloud's
geolocation API

Finding the nearest
rescue authorities
using Google Cloud's
geolocation API

SMS Alert send to
nearest authorities
using TWILIO's virtual
number

Local alert starts by
connected speakers

Generated Voice
output

Working of Nav-Voice Device

