

Hand Gesture Recognition

DATA 606: Capstone Project in Data Science

Daksh Intwala
Jaydeep Radadia
Kratika Aggarwal

Problems

The hearing impaired would benefit from greater accommodations aid in the difficulties they face.

Lack of opportunities

01

Fewer educational and job opportunities due to impaired communication.

Social Withdrawal

02

Reduced access to services and communication barrier with peers.

Emotional Issues

03

Low self esteem and confidence.

Who are our target audience?

The hearing, speaking impaired



.....

Expressing oneself

It is extremely expensive & time consuming for everyone to learn the language. This will help in live communication wherein the ASL user will be able to directly communicate with the everyone.



.....

Mass Communication

Speech generation is possible which can give voice to the ASL user. It expands the scope of communication for the conversation.



.....

Scope of Language

ASL user can communicate in multiple languages if the text output is translated in different languages using transfer learning (BERT-Hugging Face).

Literature Review

PROJECT - 01

CNN - Letter detection/recognition (93%)

Source: <https://www.kaggle.com/code/danielmarom/sign-language-using-cnn-93-with-new-images>

Methodology:

- a. Read dataset
- b. Divide it into train/test
- c. Sequential model - relu +softmax
 - i. Loss = sparse_categorical_crossentropy
 - ii. Optimizer = adam
 - iii. Metric = accuracy

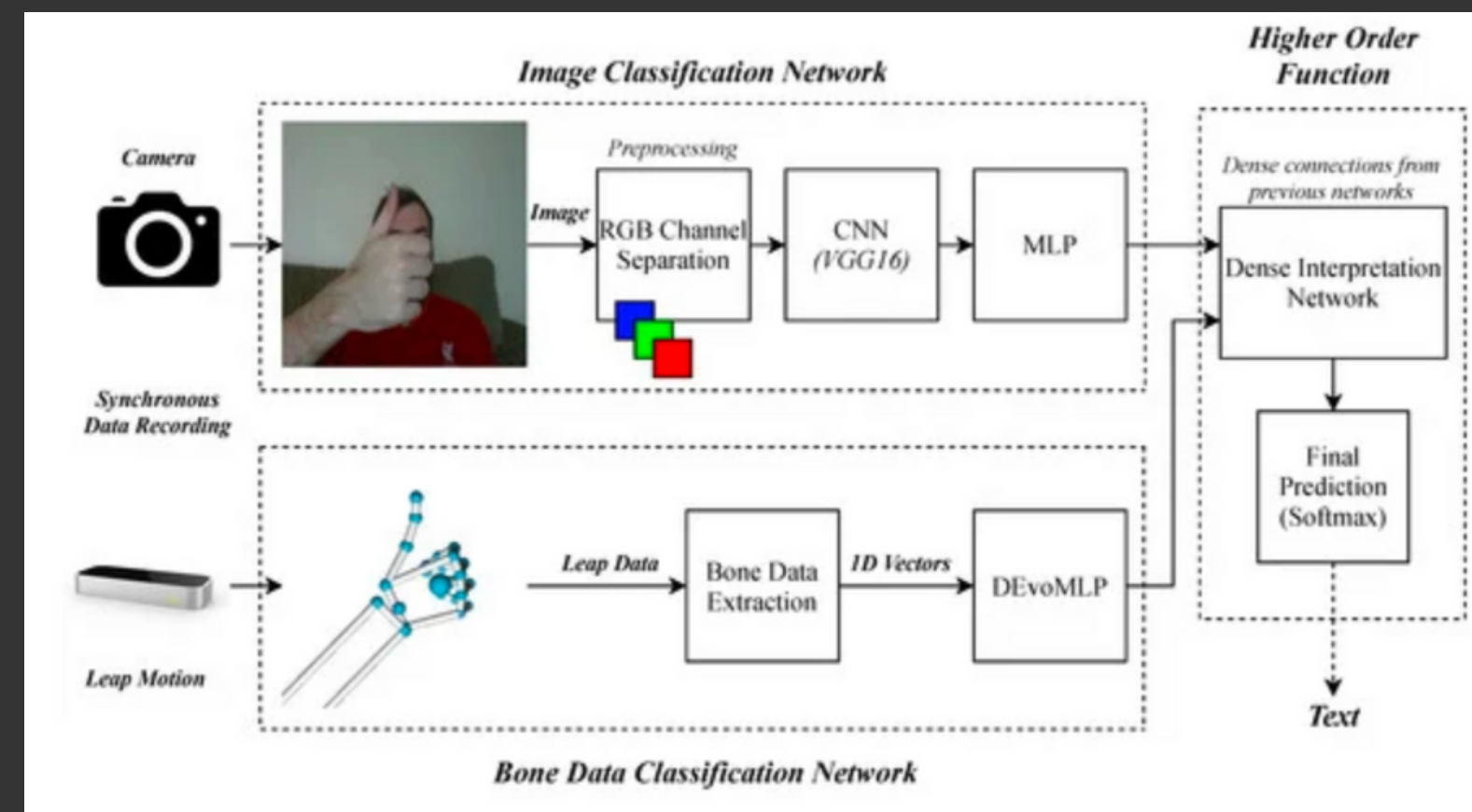
Literature Review

PROJECT - 02

British Sign Language Recognition via Late Fusion of Computer Vision and Leap Motion with Transfer Learning to American Sign Language

Source: <https://www.mdpi.com/1424-8220/20/18/5151>

Methodology:



Literature Review

PAPER - 01

Hand gestures for emergency situations: A video dataset based on words from Indian sign language

Source: <https://www.sciencedirect.com/science/article/pii/S2352340920309100>

Methodology:

In this paper, the author has analyzed the images of hand gestures indicating sign language using GoogLeNet CNN model coupled with LSTM architecture and evaluated the performance using F, Recall and Precision scores.

Literature Review

PAPER - 02

Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison

Source: https://openaccess.thecvf.com/content_WACV_2020/papers/Li_Word_level_Deep_Sign_Language_Recognition_from_Video_A_New_Large-scale_WACV_2020_paper.pdf

Methodology:

Here, two different deep learning approaches are used to detect sign language from a video i.e.,

- (i) holistic visual appearance based approach, and
- (ii) 2D human pose based approach.

Data Source

WLASL

It is the largest video dataset for Word-Level American Sign Language (ASL) recognition

It was created with the intention of facilitating the research in sign language understanding and eventually benefit the communication between deaf and hearing communities.

LINK

<https://www.kaggle.com/code/risangbaskoro/reorganize-video-data-wlasl/data>



Data



Word : Accomplish

5.4 GB

DATA VOLUME

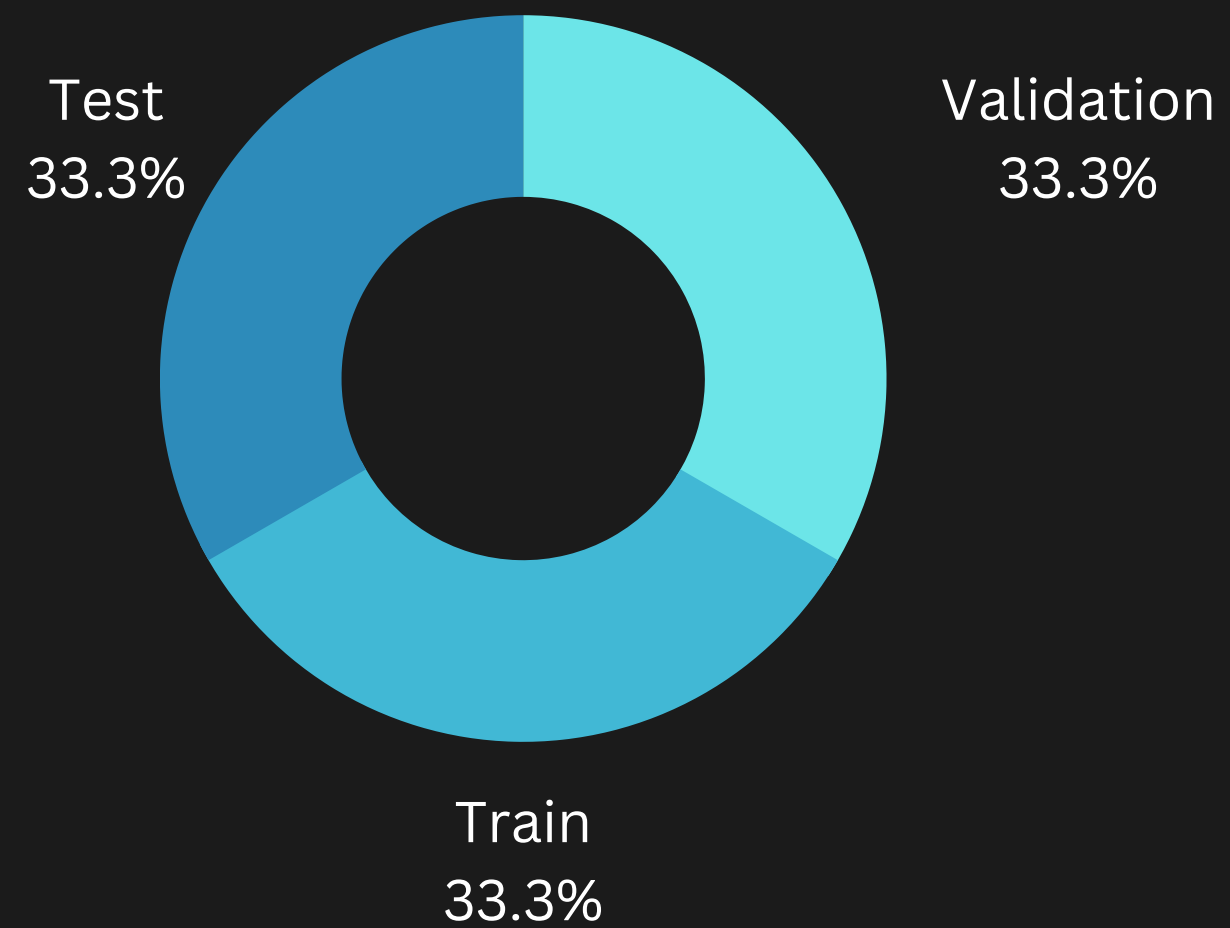
~5

CLIPS PER WORD

2000

WORDS

Data Preparation



Reorganize Video Data - WLASL

Notebook **Data** Logs Comments (0)

0 Copy & Edit 20

Data

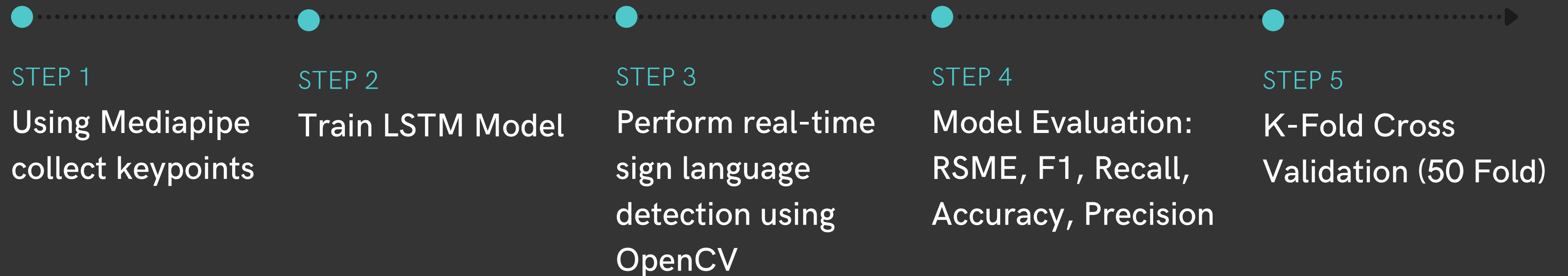
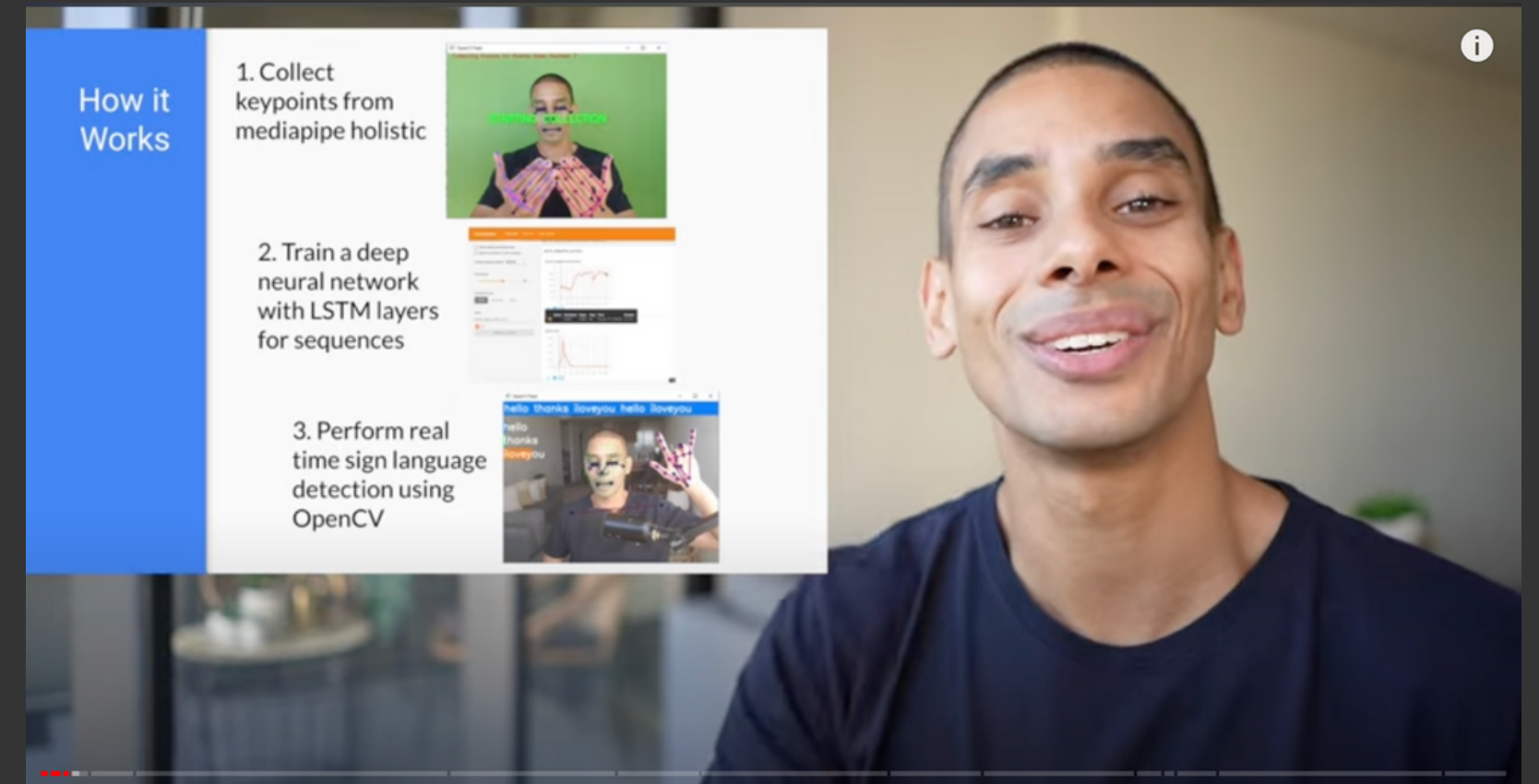
alone

02103.mp4 83.99 kB	02104.mp4 804.05 kB	02105.mp4 666.48 kB	02106.mp4 729.79 kB	02108.mp4 53.24 kB
02109.mp4 50.11 kB	02110.mp4 46.03 kB	02112.mp4 944.1 kB	65057.mp4 102.7 kB	

- airplane
- alarm
- alcohol
- algebra
- all day
- all
- allergy
- alligator
- allow
- almost
- alone**
 - 02103.mp4
 - 02104.mp4
 - 02105.mp4
 - 02106.mp4
 - 02108.mp4
 - 02109.mp4
 - 02110.mp4
 - 02112.mp4
 - 65057.mp4
- alphabet
- already
- also
- always
- amazing
- america
- amputate

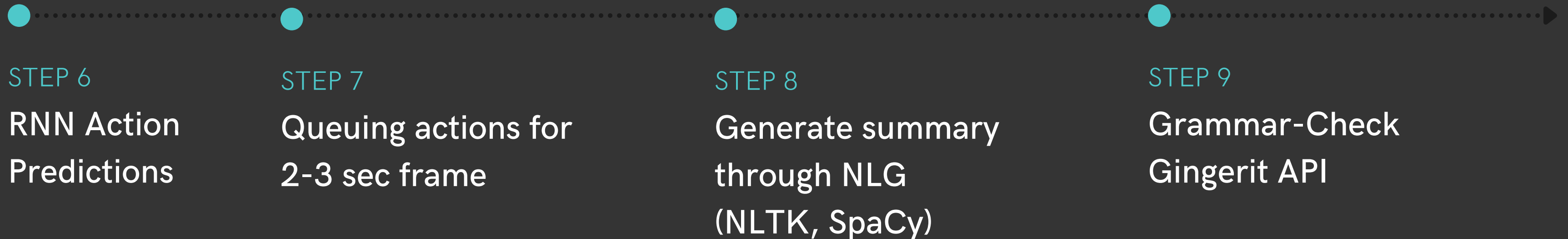
Methodology - Part A

Algorithm - Classification




Methodology - Part B

Processing & Text Generation



GitHub

https://github.com/daksh-intwala/HandGesture_Recognition_Modeling

 daksh-intwala / HandGesture_Recognition_Modeling

Private

Unwatch

1

Fork

0

Star

0

<> Code

Issues

Pull requests

Actions

Projects

Security

Insights

Settings

master


1 branch

0 tags


Go to file


Add file

<> Code

 daksh-intwala Jaydeep_Video_To_Frames

ec99431 yesterday 3 commits

 HGRM_Parsing_Clips_Frames.ipynb Jaydeep_Video_To_Frames yesterday

 README.md Clips to frames done - J. yesterday

README.md

HandGesture_Recognition_Modeling\

Aim:

We aim to develop an image classification model to perform live streaming sign language recognition which converts sign language into text format using image recognition techniques through DNN model and deliver text phrases to indicate context in the live video. To test the trained model, we plan to use the confusion matrix as a performance matrix where we would evaluate F1, Recall and Precision scores to determine model reliability and validate the results with cross validation scores using GridSearchCV. Using OpenCV to create the DNN network is our goal.

Status

1. Divided single video clip into optimal frames 0.067 setting. - Jaydeep
- 2.

About

No description, website, or topics provided.

Readme

0 stars

1 watching

0 forks

Releases

No releases published

[Create a new release](#)

Packages

No packages published

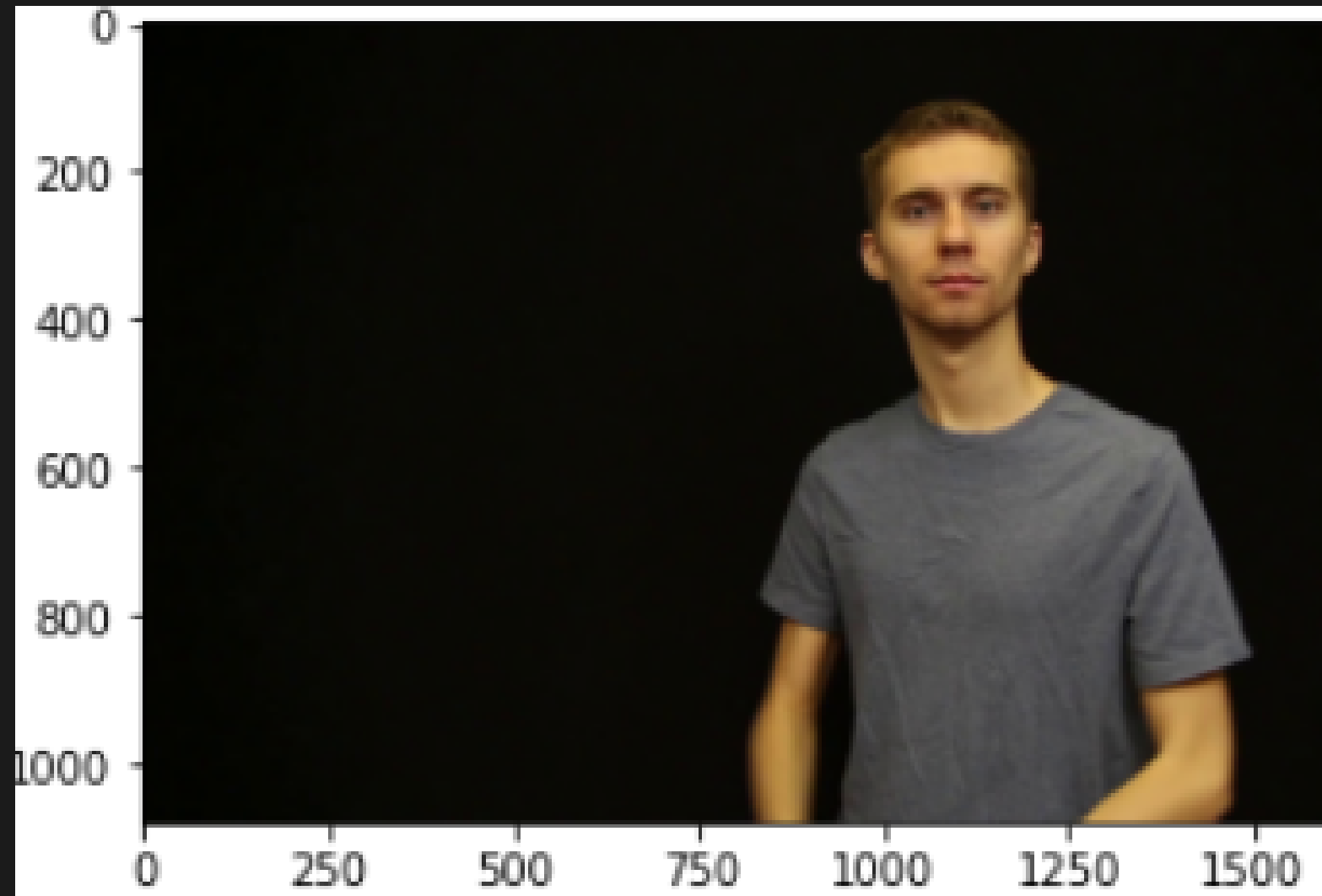
[Publish your first package](#)

Languages

Jupyter Notebook 100.0%

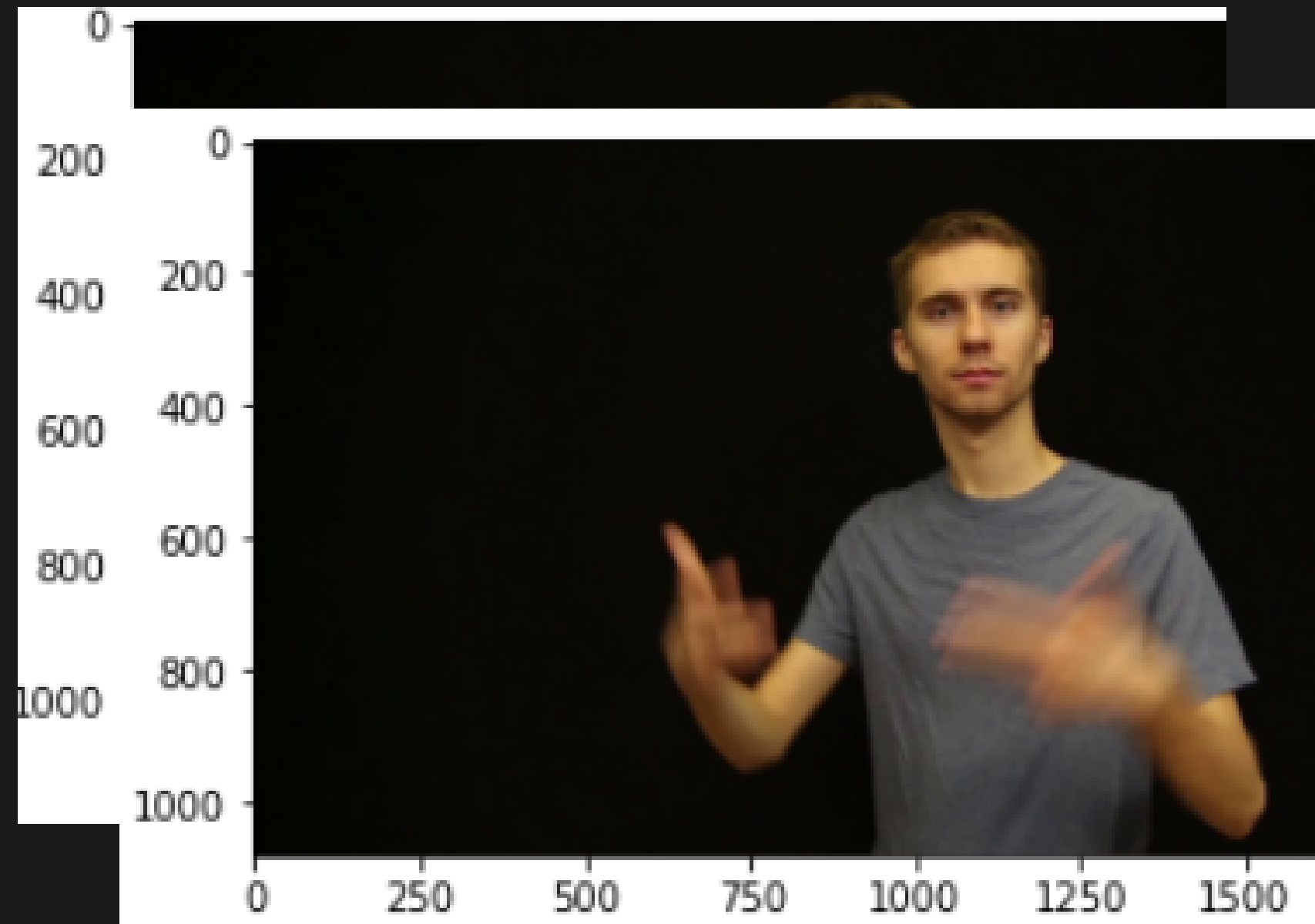
Pseudo Code

1. Parse the videoclip using openCV
2. Determine the optimum frame rate
3. Capture the frames using the .set() method
4. Save the frames



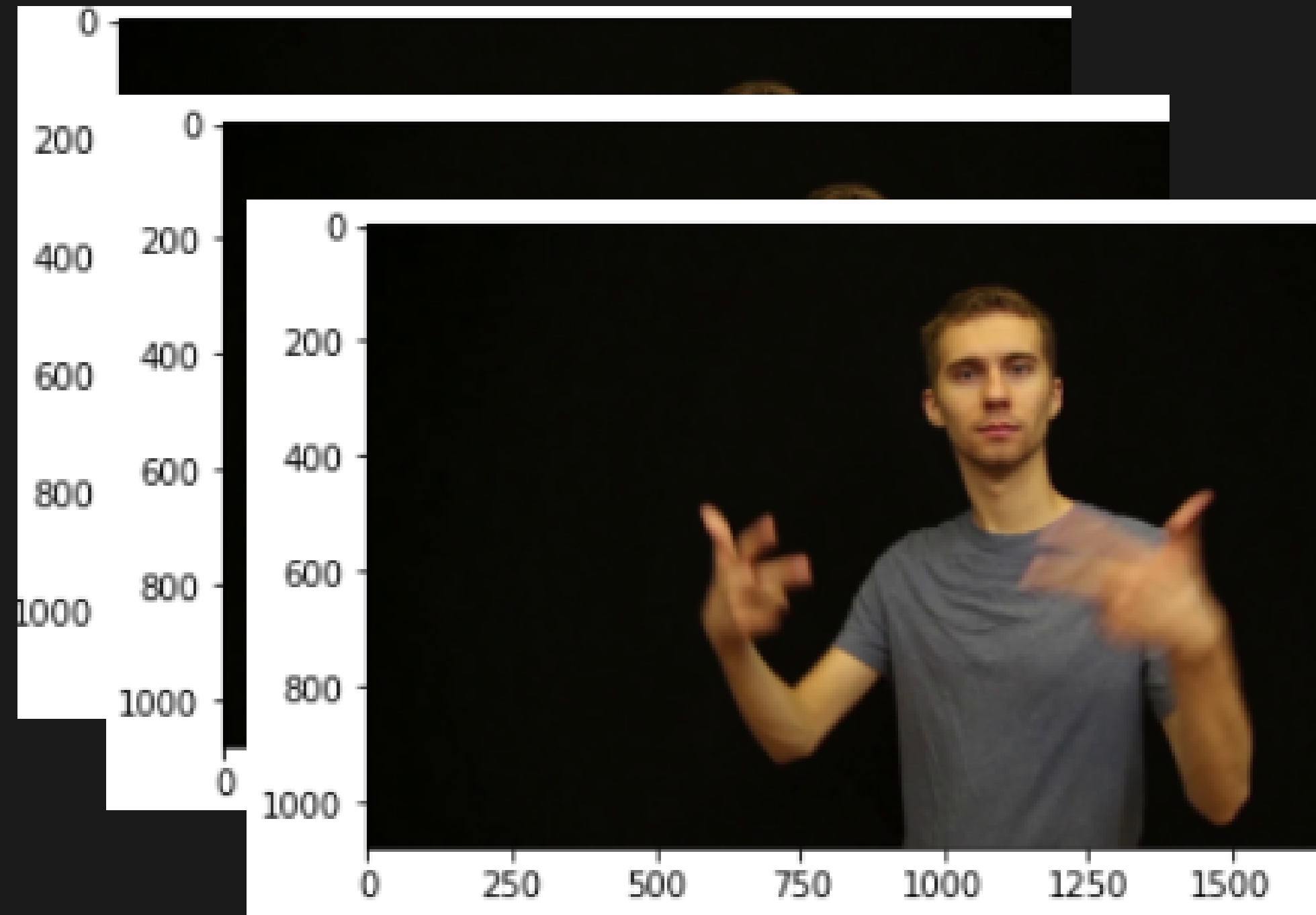
Pseudo Code

1. Parse the videoclip using openCV
2. Determine the optimum frame rate
3. Capture the frames using the .set() method
4. Save the frames



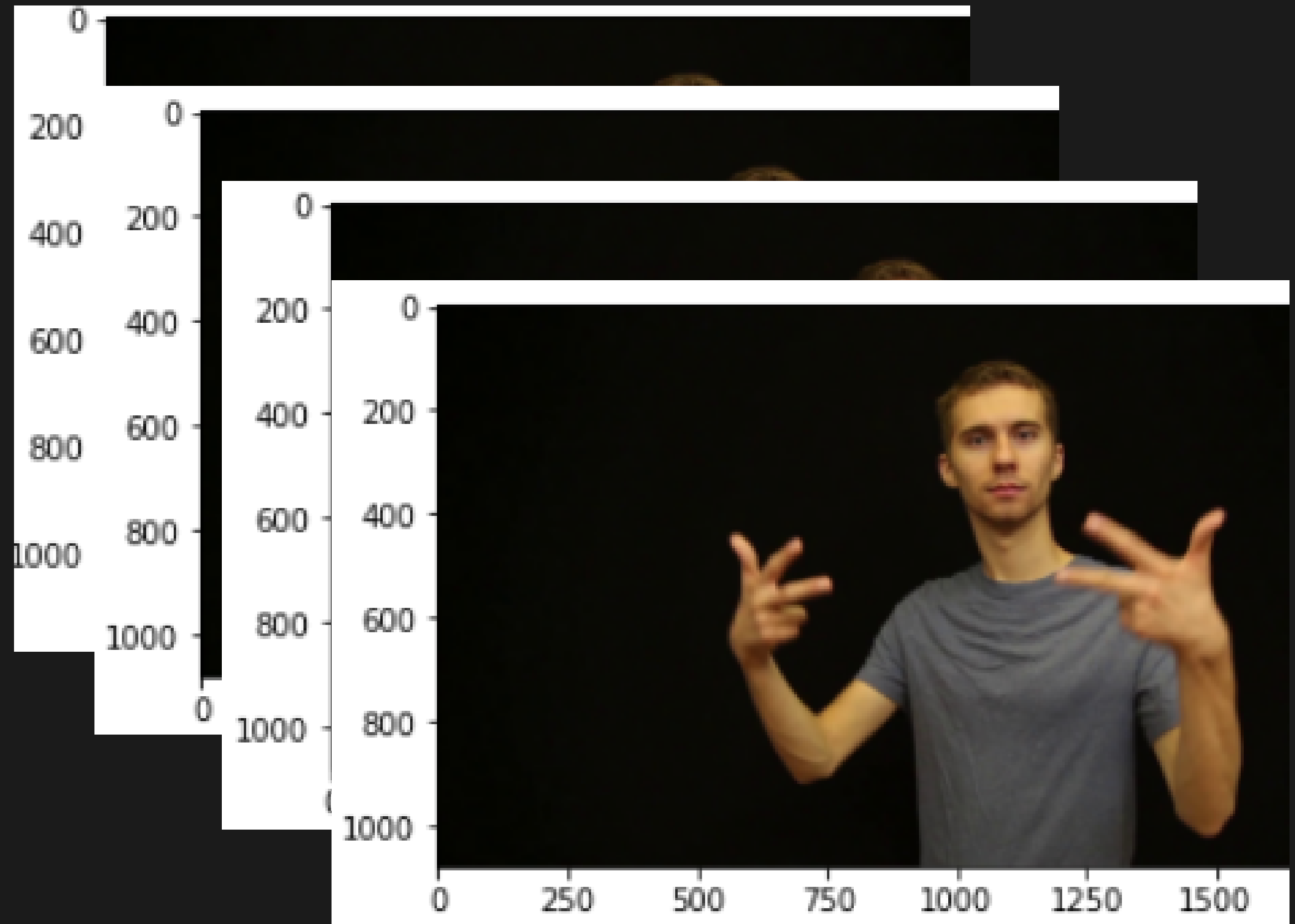
Pseudo Code

1. Parse the videoclip using openCV
2. Determine the optimum frame rate
3. Capture the frames using the .set() method
4. Save the frames



Pseudo Code

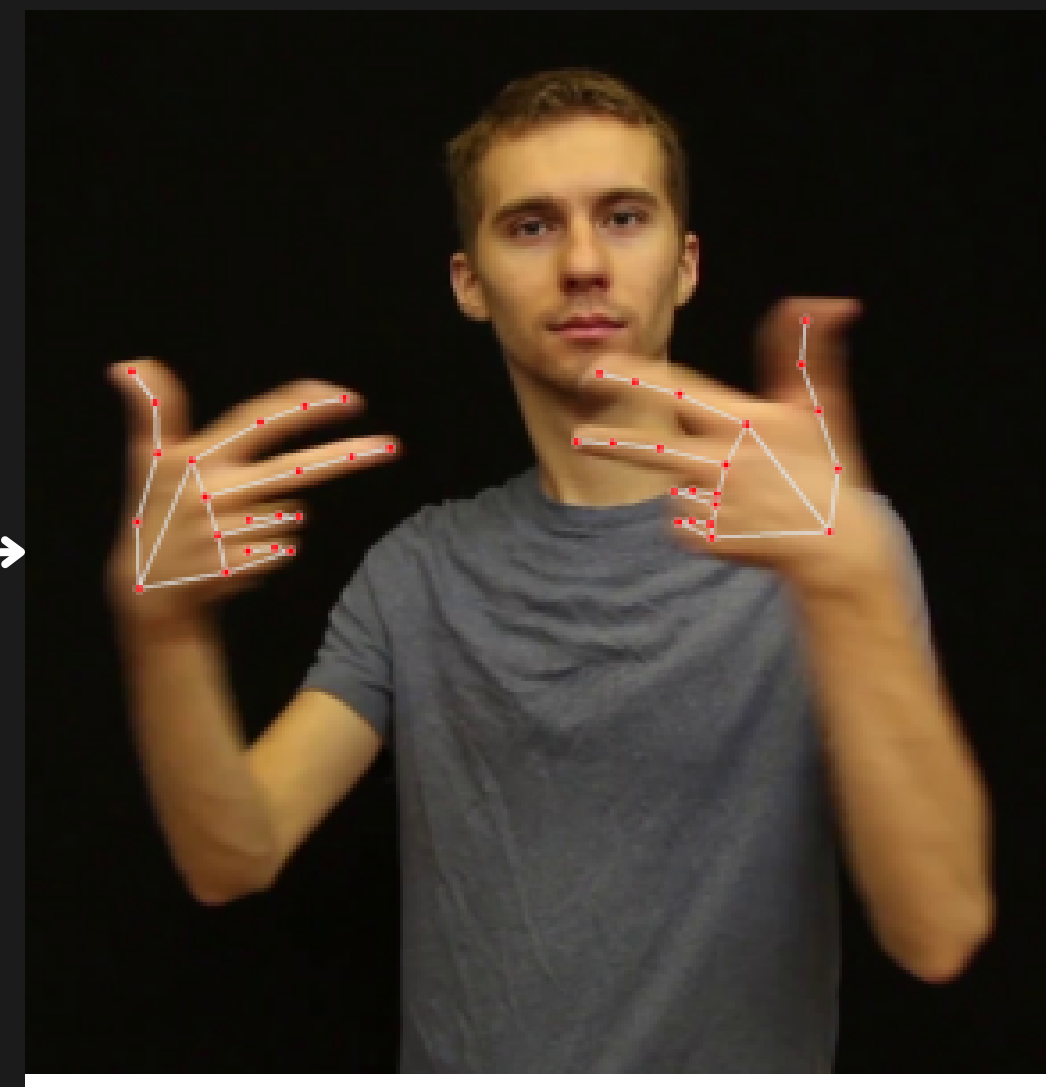
1. Parse the videoclip using openCV
2. Determine the optimum frame rate
3. Capture the frames using the .set() method
4. Save the frames



Landmarks



Hand No. 1
WRIST:
x: 667.0050430297852
y: 612.0800757408142
z: 0.0007318987263715826
THUMB_CMC:
x: 665.6389045715332
y: 548.7694072723389
z: -9.656334221363068
Hand No. 2
WRIST:
x: 1333.542366027832
y: 557.2280859947205
z: 0.0004489912407734664
THUMB_CMC:
x: 1341.1071395874023
y: 497.4174106121063
z: -29.865333437919617



References

1. <https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/hearing-loss-how-it-affects-people>