

Convert Portrait videos to vertical:

The main focus is the person class and when multiple people appear then choose according to certain criteria, this leads to facing various challenges, some main ideas to use that are needed:

- Scene detection or camera angle detection
- Person tracking or face tracking
- Face detection
- Speaker detection
- Stabilizer for the crop

The previous method are to overcome challenges like:

- Detect people and if only one appears focus on him
- If two people then use the face to choose, if only 1 face choose the person of the face if more than one face detected choose the speaker, if more than one is speaking you can choose the second speaker
- Looking at the future data may be important like what will happen may affect what is happening
- The crop should be smoothly moving in case focusing on moving person and if the face of a person does not get out of the crop there is no need to move the crop coordinates
- If someone is speaking for a long time switch on the other person for few seconds so that we know he is talking to somebody (if detecting reaction is easy we can use it to show the other person)
- Always check for overlap between people so that if multiple faces can appear in a crop choose to center the crop between them but if there are for example 3 people and one in the middle is overlap with both people on the edges then choose the middle overlap in one of the edges