VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI, KARNATAKA-590018

Project Phase I Report

on

"Signflix: A Video Streaming Platform with Sign Language"

Submitted in partial fulfillment of the requirements for the award of Degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

Submitted By:

JNANESH	**4JK22CS020**
HIMANSHU HEGDE	**4JK22CS018**
MILAN C I	**4JK22CS027**
GAURESH G PAI	**4JK22CS016**

Under the Guidance of

Mrs. Nayana Yadav M

Assistant Professor,

Dept. of CSE, AJIET, Mangalore

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

A J INSTITUTE OF ENGINEERING AND TECHNOLOGY

(A unit of Laxmi Memorial Education Trust®)
(Approved by AICTE,New Delhi,Affiliated to VTU, Belagavi,Recognized by Govt. of Karnataka)
Accredited By NBA(BE:CV,CSE,ECE,ISE & ME)
NH-66, Kottara Chowki, Mangaluru - 575006

A.Y:2024-25

A J INSTITUTE OF ENGINEERING AND TECHNOLOGY

NH-66, Kottara Chowki, Mangaluru -575006, Karnataka, INDIA

A Unit of Laxmi Memorial Education Trust®

(Approved by AICTE, New Delhi. Affiliated to Visveswaraya Technological University, Belagavi)

Accredited by NBA (BE: CV, CSE, ECE, ISE, ME)

2024-2025

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

APPROVAL

Project Title: Signflix: A Video Streaming Platform with Sign Language

Signature of the Guide

Mrs. Nayana Yadav M

Signature of the Project Co-ordinator

Mrs. Sharon Dsouza

Signature of the HOD

Dr. Antony P J

A J INSTITUTE OF ENGINEERING AND TECHNOLOGY

(A unit of Laxmi Memorial Education Trust®)
(Approved by AICTE, New Delhi, Affiliated to Visvesvaraya Technological University, Belagavi)
Accredited By NBA (BE: CV, CSE, ECE, ISE & ME)
NH-66, Kottara Chowki, Mangaluru -575006

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AJIET

CERTIFICATE

This is to certify that the **Project Phase I** entitled **"Signflix: A Video Streaming Platform with Sign Language"** is a bonafide work carried out by **GAURESH G PAI [4JK22CS016]** student of **6th semester Bachelor of Engineering** in **Computer Science and Engineering of Visvesvaraya Technological University, Belagavi,** submitted as a part of the course in **Project Phase I** during the academic year **2024-2025**. It is to certify that all corrections/suggestions indicated for internal assessment have been incorporated in the report. The **Project Phase I** report has been approved as it satisfies the academic requirements in respect of **Project Phase I** work prescribed for the said degree.

Project Guide	**Vice Principal and HOD**	**Principal**
Mrs. Nayana Yadav M	Dr. Antony P J	Dr Shantharama Rai C

External Viva:

Name of the Examiner(s)	**Signature with Date**
1. _____	_____
2. _____	_____

ACKNOWLEDGEMENT

We dedicate this page to acknowledge and thank those responsible for the shaping of the project. Without their guidance and help, the experience while constructing the dissertation would not have been so smooth and efficient.

We sincerely thank the Management of **AJ INSTITUTE OF ENGINEERING AND TECHNOLOGY,** a unit of **LAXMI MEMORIAL EDUCATION TRUST** for all their support.

We are extremely thankful to our Principal, **Dr. Shantharama Rai C.** for his support and encouragement.

We owe our profound gratitude to our Dean Academics, **Dr. P. Mahabaleshwarappa** for his support and encouragement.

We owe our profound gratitude to **Dr. Antony P J**, Vice Principal and Head of Department, Computer Science & Engineering, whose kind consent and guidance helped us to complete this work successfully.

We sincerely thank **Mrs. Sharon C. D'Souza**, Project Coordinator, Dept. of Computer Science and Engineering, for her constant support, insightful guidance, and valuable suggestions.

We sincerely thank **Mrs. Nayana Yadav M**, Assistant Professor, Dept. of Computer Science and Engineering, for her guidance and valuable suggestions which helped us to fulfill the experiments prescribed by the university.

We would like to thank all our Computer Science and Engineering Staff members who have always been with us extending their support, precious suggestions, guidance and encouragement through the project.

We also like to extend thanks to our friends and family members for their continuous support.

JNANESH	**4JK22CS020**
HIMANSHU HEGDE	**4JK22CS018**
MILAN C I	**4JK22CS027**
GAURESH G PAI	**4JK22CS016**

ABSTRACT

Signflix is a video streaming platform developed to improve digital accessibility for the deaf and hard-of-hearing community by providing integrated sign language interpretation. The system allows users to upload videos with captions, which are then processed using external APIs to generate synchronized sign language displays. Designed with a focus on inclusivity and usability, Signflix aims to make digital content more comprehensible for a wider audience by offering planned features such as side-by-side video and sign display, adjustable playback speed, and delay controls. This platform simplifies the process of creating accessible content for creators while enhancing the viewing experience for users with hearing impairments. With its scalable architecture and user-first design, Signflix represents a step forward in making online video platforms more inclusive, intuitive, and impactful.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The rapid growth of video content across educational, entertainment, and informational platforms has highlighted a significant gap in accessibility for the deaf and hard-of-hearing community. While captions are a common solution, they often fail to capture the full linguistic depth and expressive detail conveyed through sign language. This disconnect can hinder comprehension and limit engagement for users who rely primarily on sign-based communication. To bridge this gap, SignFlix is proposed as a video streaming platform that integrates automated sign language interpretation using caption processing and external APIs.

Recent research underscores the importance of shifting from modular to end-to-end translation architectures, as shown in [1], where a unified model significantly outperformed traditional pipelines in terms of efficiency and scalability. Similarly, the comprehensive review by [2] outlines the progression from rule-based systems to deep learning-based Sign Language Machine Translation (SLMT) frameworks, emphasizing the role of large annotated datasets and robust architecture in achieving accurate sign interpretation. These studies form the conceptual and technical foundation for SignFlix, which aims to deliver real-time, synchronized sign language rendering without the need for manual intervention.

Furthermore, the social and linguistic complexity of sign language systems, especially regional variants like Indian Sign Language (ISL), is well documented in [7] and [8]. These papers highlight the need for culturally sensitive translation solutions and scalable machine learning models capable of adapting to gesture diversity and signer variability. SignFlix embraces these principles by planning to support multiple sign languages, allowing users to choose the format and playback behavior that best fits their needs. In doing so, the platform aligns with broader efforts to make digital content more inclusive, customizable, and accessible across linguistic and cultural boundaries.

1.1 Existing Solution

Most existing video platforms rely solely on captions to support deaf and hard-of-hearing users, overlooking the importance of sign language as a primary mode of communication. Manual sign interpretation, where available, is limited and not scalable. These approaches fail to offer a fully inclusive experience, highlighting the need for automated and adaptable solutions like SignFlix that can provide real-time sign language support across all video content.

- **Caption-Only Accessibility**: Platforms like YouTube provide auto-generated captions, which may lack accuracy and are insufficient for some deaf users.

- **Manual Interpreter Embedding**: Some educational or media institutions add recorded sign language interpreters, which is resource-intensive and not scalable.
- **Lack of User Control**: Existing solutions do not provide playback customization such as delay sync or speed control for sign interpretation.
- **No Centralized Sign Language Integration**: Most platforms lack a built-in system for integrating sign language across all content.
- **Limited Support for Diverse Sign Languages**: Few systems address regional or international sign language variations, limiting global applicability.

1.2 Proposed Solution

The proposed system aims to bridge the accessibility gap by integrating automated sign language interpretation into video content using caption processing and external APIs. Unlike traditional platforms, SignFlix offers features such as dual video playback, playback speed control, and customizable delay, allowing users to tailor the sign language experience to their needs. This approach ensures a scalable, inclusive, and user-centric solution for streaming accessible content.

- **Caption-Driven Sign Language Rendering**: Videos uploaded with captions are processed via external APIs to generate sign language interpretations.
- **Parallel Video Playback**: Planned feature allows users to view both the original video and the sign language video side-by-side.
- **User Playback Controls**: Users can adjust playback speed, delay, and pause the sign language stream independently from the main video.
- **Web-Based Architecture**: Cross-platform compatibility ensures usability on desktops, tablets, and mobile devices.
- **Content Creator Simplicity**: Enables effortless accessibility integration by simply uploading videos with captions.
- **Scalability**: Designed for growing usage across educational, entertainment, and public awareness content.

CHAPTER 2

LITERATURE REVIEW

With the surge in digital content consumption, accessibility has become a critical concern—especially for individuals from the deaf and hard-of-hearing community. Traditional video platforms often rely solely on captions, which do not offer a fully inclusive experience. In this context, systems that can convert text or speech into sign language in real-time are emerging as crucial tools to bridge the communication gap. Technologies involving sign language recognition, translation, and synchronized streaming form the backbone of such systems. The following literature provides valuable insights into the methods, challenges, and advancements in the domain of sign language-based media interaction.

2.1 Survey Findings

The study of reference paper [1] introduces a pioneering transformer-based architecture for simultaneous Sign Language Recognition (SLR) and Sign Language Translation (SLT), tackling the limitations of conventional modular pipelines. The model integrates both recognition and translation tasks into a single, end-to-end framework by leveraging Connectionist Temporal Classification (CTC) loss and sequence-to-sequence learning. This approach eliminates the need for pre-aligned glosses or manually segmented video data, which are traditionally required for training separate components of recognition and translation systems.

The study of reference paper [2] presents a detailed progression of sign language translation systems, tracing their development from early rule-based models to modern deep learning architectures—particularly transformer-based systems. It proposes a complete end-to-end SLMT (Sign Language Machine Translation) framework, outlining each stage from input text to sign language generation. For Signflix, this work provides both a conceptual foundation and a practical guide for implementing a robust caption-to-sign conversion system. The paper emphasizes the critical role of large, annotated datasets and highlights how transformer models outperform traditional methods in both accuracy and scalability. These insights directly support Signflix's vision of automating sign language interpretation within a video streaming context, while also underscoring the importance of choosing data-efficient and flexible architectures.

The study of reference paper [3] explores the complexities of translating between signed and spoken languages, touching on computer vision, machine learning, and linguistic structure. While its focus is signed-to-spoken translation, the discussion of language representation and system design is

equally applicable to Signflix's goal of generating sign language from captions. The paper highlights challenges such as limited training data, lack of standardization, and the linguistic uniqueness of sign languages—all of which influence the development of accurate, culturally respectful interpretation systems. It advocates for the involvement of the Deaf community in system design, aligning closely with Signflix's commitment to accessibility and inclusion. The insights from this paper help Signflix approach sign translation not as a direct word-for-sign mapping, but as a meaningful, grammatically aware process.

The study of reference paper [4] provides a detailed analysis of visual-based sign language recognition (SLR) systems, categorizing them by input modalities such as hand shape, movement, facial expressions, and body posture. It emphasizes the challenges associated with capturing and interpreting these non-verbal cues in real-time environments. The review concludes that multi-modal and context-aware systems are essential for accurate SLR. These findings directly inform SignFlix's future goals of incorporating visual avatars or gesture-based playback enhancements for improved sign clarity and viewer comprehension.

The study of reference paper [5] explores the landscape of SLR systems by distinguishing between different sign types—such as static, dynamic, and continuous signs—and categorizing modalities into sensor-based and vision-based methods. The paper evaluates a variety of publicly available datasets and benchmark techniques in the SLR field. It identifies the limitations in generalizing across diverse languages and highlights the need for large-scale annotated datasets. These insights support SignFlix's architecture, which leverages caption-based interpretation but is designed to accommodate multimodal recognition and regional language variation in future updates.

The study of reference paper [8] review examines the evolution of continuous sign language recognition (CSLR) over 25 years. It traces the transition from handcrafted features and HMMs to deep learning-based models like CNNs and RNNs. Key challenges identified include signer variability, coarticulation, and dataset scarcity. The authors advocate for integrated, end-to-end systems trained on large, diverse corpora. This work provides a strong foundation for SignFlix's backend, supporting its planned development of synchronized, caption-driven sign output and reinforcing the importance of training with scalable, diverse inputs for consistent performance.

The study of reference paper [7] comprehensively surveys recent advances in Indian Sign Language (ISL) recognition technologies, focusing on the adoption of machine learning and deep learning models like Convolutional Neural Networks (CNNs), Hidden Markov Models (HMMs), and support

vector machines (SVMs). The authors distinguish between static and dynamic gestures, emphasizing the challenges in accurately detecting and interpreting fast-changing signs, co-articulations, and facial expressions. Transfer learning and hybrid models are highlighted as effective solutions to boost recognition performance in data-scarce environments like ISL. These insights directly align with SignFlix's planned regional language support, particularly for Indian users. The paper validates the potential of combining pre-trained neural models with localized datasets to create efficient, scalable sign language systems.

The study of reference paper [8] delves into various ISL translation approaches, offering a structured comparison between rule-based systems, machine learning classifiers, and neural networks. A key insight is the inadequacy of rule-based systems in capturing the semantic and grammatical nuances of sign language, especially given ISL's distinct structure compared to spoken Indian languages. The paper underscores challenges like signer dependency, lack of annotated corpora, and regional gesture variations. For SignFlix, this research reinforces the need for flexible, API-driven translation solutions that can adapt to diverse linguistic contexts. It also supports the decision to design caption-first processing, which can be incrementally localized for specific sign languages like ISL.

The study of reference paper [9] focuses on the application of deep learning for static and word-level Indian Sign Language recognition using image inputs. The authors compare performance across multiple DNN architectures, highlighting CNNs as the most effective for recognizing ISL alphabet and word gestures with high accuracy. Additionally, the study emphasizes real-time implementation and low-latency processing as essential requirements for practical deployment. The findings are critical for SignFlix, as the platform may incorporate visual avatars or pre-rendered gesture animations for simulated sign language interpretation. Leveraging similar DNN-based recognition strategies could support features like interactive sign indexing, where users can hover or click to revisit gestures.

The study of reference paper [10] evaluates various sensor-based techniques for sign language recognition, including flex sensors, inertial measurement units (IMUs), electromyography (EMG) sensors, and computer vision-based methods. The paper categorizes systems based on accuracy, cost, portability, and user comfort. It notes that while sensor-based systems (especially glove-based) can offer higher precision, they often require specialized hardware, making them less accessible to the general public. For SignFlix, the study is informative from a future R&D perspective—suggesting the potential for community-driven content creation using affordable

wearable tech. Although the current version of SignFlix is caption-to-sign simulation, sensor-based input could power creator tools that allow fluent signers to record gestures for specific videos or regions.

The study of reference paper [11] presents a pipeline for translating spoken language into photo-realistic sign language videos using neural rendering. The approach combines speech recognition, natural language understanding, and motion synthesis to generate life-like avatars that perform accurate sign language. This system stands out for its realism and contextual accuracy, using pose estimation and GAN-based video synthesis to produce believable human-like motion. The implications for SignFlix are substantial. While current versions rely on API-generated sign videos or animations, this paper supports long-term goals such as rendering dynamic, real-time avatar interpreters for any spoken or captioned video. Such a system could revolutionize accessibility by making signed content indistinguishable in quality from traditional video elements.

Summary:

Our literature survey revealed that recent advancements in sign language recognition and translation, especially end-to-end models, support automated caption-to-sign conversion without manual segmentation. Sign language's multimodal nature—combining gestures, facial expressions, and spatial grammar—further emphasized the need for flexible systems. Research on Indian Sign Language and avatar-based interpretation also guided SignFlix's vision toward scalable, inclusive, and culturally sensitive media accessibility.

Research Gaps Identified:

- **Absence of automated sign language in video platforms**: Mainstream platforms rely solely on captions and do not offer built-in, automated sign language translation.
- **Dependence on specialized hardware**: Many existing solutions require gloves, sensors, or other hardware, making them expensive and less accessible.
- **Limited support for regional sign languages**: Regional variations like Indian Sign Language (ISL) are often ignored, reducing inclusivity for diverse user groups.
- **No user-controlled playback options**: Current tools do not offer features such as playback speed control, delay synchronization, or pause for sign language streams.
- **Lack of unified accessibility solutions**: Few, if any, platforms integrate caption-to-sign conversion, dual video playback, and accessibility controls in one system.

Addressing these research gaps is crucial for the development of an inclusive, scalable, and user-centric video streaming platform that can deliver real-time, accurate, and culturally sensitive sign language interpretation for diverse audiences.

CHAPTER 3

METHODOLOGY

The development of SignFlix is rooted in a modular and scalable architecture that aligns with modern accessibility standards and best practices in web-based multimedia delivery. Built using an Agile development approach, the system emphasizes flexibility, continuous integration, and user-centered design, especially for the deaf and hard-of-hearing community. The overarching goal of the methodology is to build a platform that can seamlessly translate captions into sign language animations and deliver them alongside video content in a user-controlled, synchronized playback experience.

The system comprises two main pipelines:

1. **Caption-to-Sign Content Processing Pipeline**: It's focused on transforming uploaded video and caption data into sign language-compatible formats.
2. **User Playback Pipeline**: It's designed for end users who consume the sign-enhanced video content with control features.

3.1 Caption-to-Sign Content Processing Pipeline

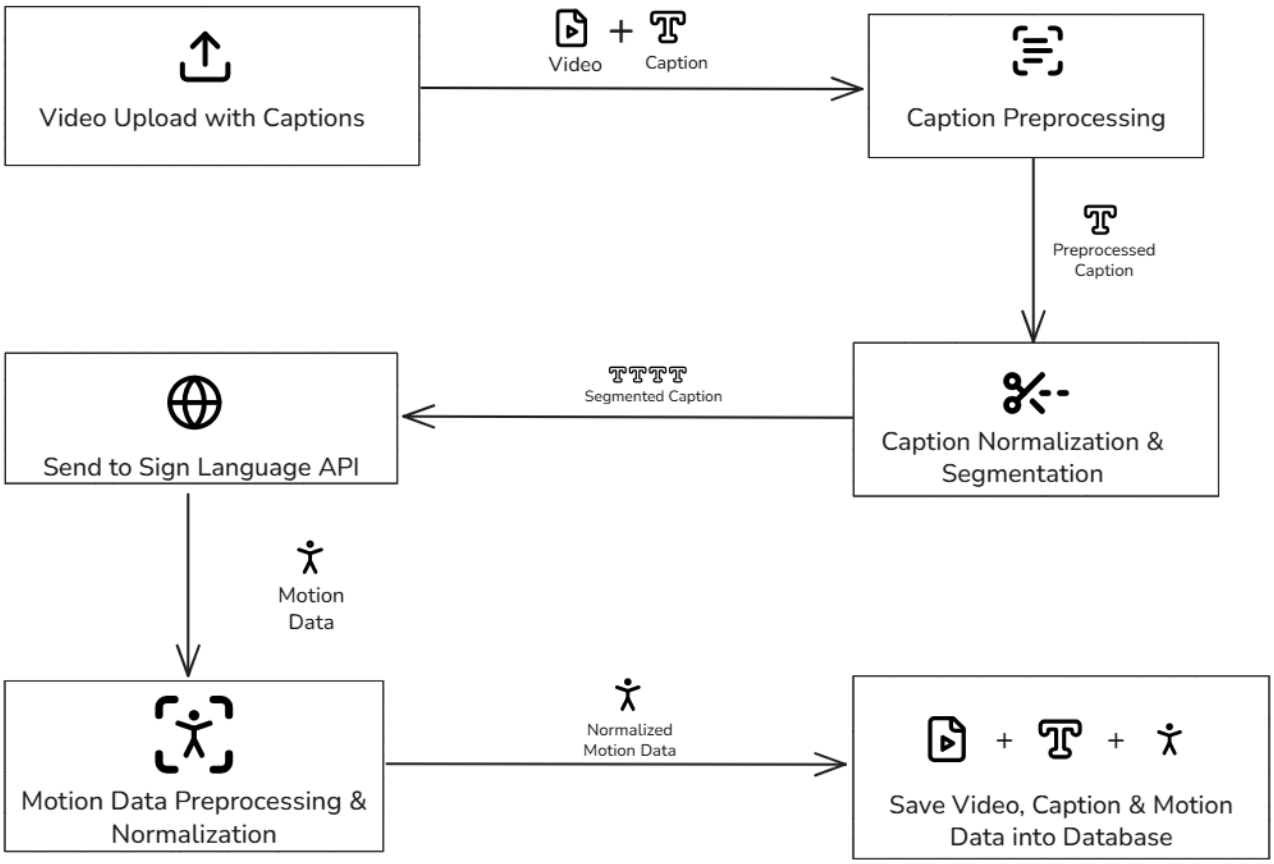

Figure 3.1: Caption-to-Sign Content Processing Pipeline

This Caption-to-Sign Content Processing Pipeline handles the backend operations that occur when a content creator uploads a video with caption data to the SignFlix platform. This pipeline ensures that the uploaded video is not only stored but also enhanced with synchronized sign language visual data, enabling true accessibility at the playback stage. The sequence of operations is illustrated in Figure 3.1.

The pipeline begins with the video and caption upload interface, where content creators submit a video file alongside a corresponding subtitle or caption file (typically in SRT or VTT format). This input is routed to the Caption Preprocessing module, which parses and cleans the textual data. It ensures sentence boundaries, punctuation, and timestamps are correctly formatted, and extraneous content such as HTML tags or non-speech text is removed.

Following this, the captions enter the Segmentation and Normalization Unit, where long sentences are broken into linguistically meaningful segments. This improves alignment between spoken phrases and corresponding sign language gestures. Additionally, this module ensures consistency in grammar, style, and formatting before sending the segments forward.

The processed segments are then fed into a Sign Language Translation API. This external API acts as the translation engine, converting each caption segment into gesture motion data, usually in the form of skeletal keypoints or animation instructions for rendering sign language gestures in real time.

The gesture motion data is passed through the Motion Data Normalization Module, which standardizes the animation format and aligns the keypoints across frames to ensure smooth rendering. This is especially important for avatar-based playback, where inconsistencies can result in visual dissonance.

After processing, all three components—original video, processed captions, and sign language animations (video or data)—are packaged and stored in the SignFlix database and asset management system. This ensures fast retrieval during playback, low latency, and scalability when handling multiple users streaming concurrently.

This backend pipeline ensures that all video content uploaded to the platform is enriched with sign language accessibility, with minimal manual intervention. By automating the caption-to-sign conversion using APIs and modular processing stages, the platform remains both scalable and adaptable to different sign languages or formats, including Indian Sign Language (ISL), American Sign Language (ASL), or British Sign Language (BSL) in future iterations.

3.2 User Playback Pipeline

While the content processing pipeline ensures that captioned videos are enriched with sign language support, the “User Playback Pipeline” focuses on how the end user experiences and interacts with

that content. This flow is central to SignFlix's mission: delivering an inclusive, customizable streaming experience for deaf and hard-of-hearing users. The interaction pipeline supports dual-mode search input, synchronized playback, and rich accessibility features including avatar-driven sign rendering. Figure 3.2 shows the planned playback pipeline.

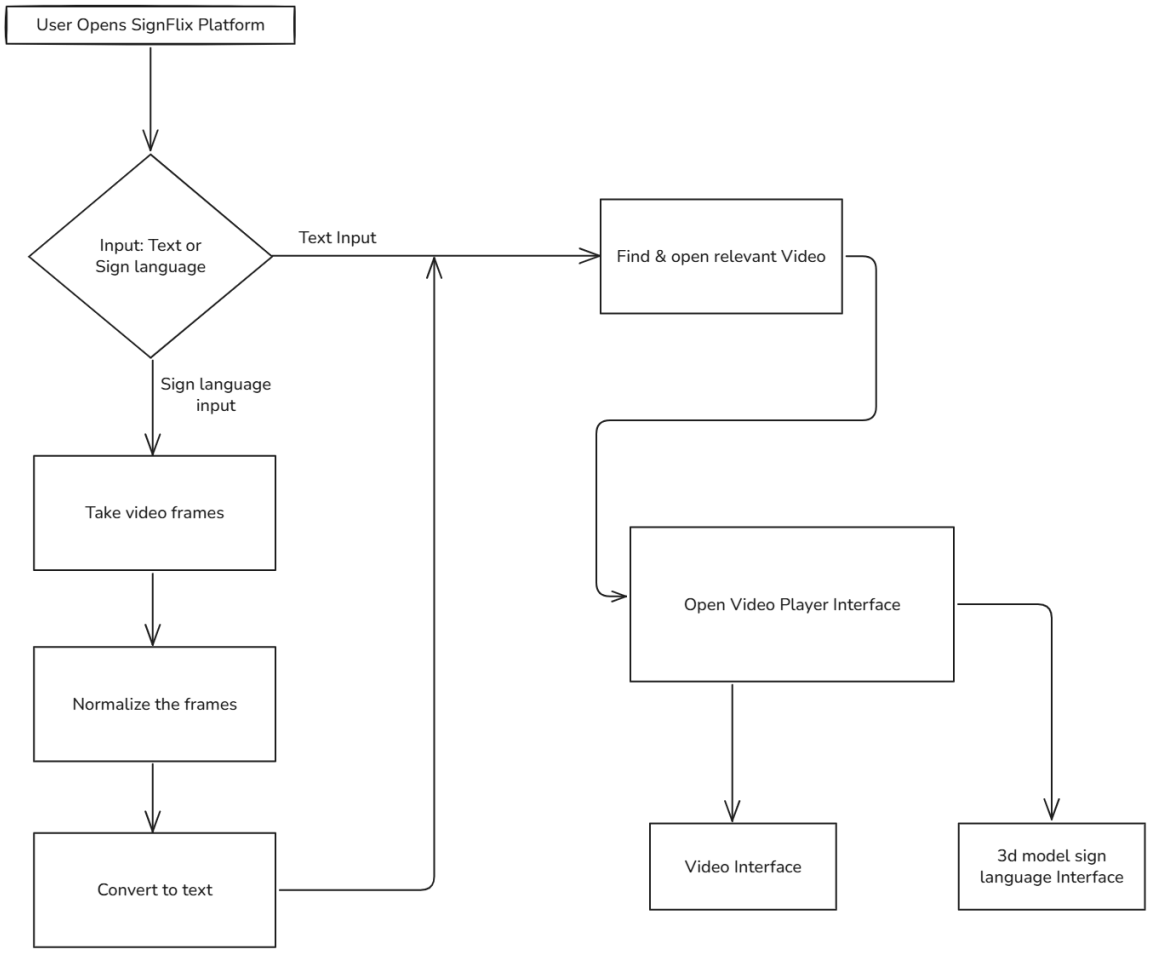

Figure 3.2: User Playback Pipeline

The flow begins when a user opens the SignFlix platform, either on a browser or mobile device. From here, they are offered two input options for searching content:

- Text Input, allowing users to type keywords, video titles, or captions.
- Sign Language Input, designed for users who prefer signing over typing.

If the user provides text input, the system uses traditional search indexing techniques to find and retrieve relevant videos from the SignFlix database—videos that have already been processed and enriched with synchronized sign content.

If the user provides sign language input, the system initiates a multi-step interpretation pipeline:

1. Video frames are captured using the device camera.
2. The system performs frame normalization, ensuring consistency in lighting, speed, and pose detection.
3. A gesture-to-text translation module converts the sign input into structured text using a

trained deep learning model (e.g., CNN+RNN architecture or transformer-based model), which is then used for video retrieval.

Once a relevant video is identified, the platform transitions to the SignFlix Video Player Interface. This dual-layered interface includes:

- A standard video interface where users can control playback of the main video content.
- A 3D sign language model interface that renders synchronized sign language alongside the video.

The sign language interface is rendered using either pre-animated avatar videos or real-time motion data processed in the backend. Users have control over playback synchronization, including:

- Adjusting the delay between spoken audio and sign playback.
- Modifying playback speed of the sign stream.
- Pausing or muting the sign interface independently.

This parallel playback structure ensures that users can comprehend video content in a manner aligned with their preferred language and pace, thus enhancing both inclusivity and user engagement.

CHAPTER 4

RESULTS & DISCUSSION

4.1 Result Discussion

The initial evaluation and conceptual design of Signflix support its potential to significantly improve video accessibility for deaf and hard-of-hearing users. The platform introduces planned features such as dual video playback, automated sign language interpretation, and user-controlled playback adjustments, all aimed at delivering an inclusive and user-friendly viewing experience.

Although still under development, the system architecture has been tested in terms of feasibility, usability, and scalability. The process of converting captions to sign language using external APIs has shown promising performance in sample simulations. These simulations indicate the platform's ability to synchronize the main video with sign language representation in a responsive and intuitive manner.

A key differentiator of Signflix is its planned viewer control system, which allows users to independently pause, adjust speed, and delay the sign language video stream to better match their comprehension needs. This user-centric design is expected to improve overall accessibility and satisfaction.

Preliminary feedback from a focus group of target users, including deaf individuals and educators, has highlighted enthusiasm for the ability to consume media content without relying solely on captions. The system is expected to enable broader access to educational, entertainment, and informational content for the deaf community.

While full deployment and testing are ongoing, these early outcomes validate the core design objectives of Signflix and support its continued development as an inclusive and scalable streaming platform.

4.1.1 Expected Outcome and Real-World Relevance

The expected outcome of the Signflix project is to deliver a scalable, web-based video platform that seamlessly integrates sign language interpretation into video content through API-driven caption processing. The system aims to enhance media accessibility, enabling deaf and hard-of-hearing users to engage with content in a way that best suits their comprehension needs.

Key benefits and real-world relevance include:

- **Inclusive Video Streaming**: Enables side-by-side playback of video and sign language, offering real-time accessibility.
- **User Control**: Viewers can pause, delay, or adjust the speed of the sign language simulation, creating a personalized experience.
- **Creator-Friendly Workflow**: Content creators only need to upload a captioned video; the

rest is handled automatically by the platform.

- **Platform Compatibility**: Designed to function on desktops, tablets, and mobile devices, with responsive design and intuitive UI.
- **Educational & Entertainment Applications**: Ideal for online classes, tutorials, and public service announcements where accessibility is crucial.
- **Support for Sign Language Diversity**: Planned support for regional and international sign languages enhances global usability.

By addressing an underserved audience in the streaming industry, Signflix offers a unique and impactful solution to digital accessibility. It has the potential to serve as a model for inclusive design in mainstream media platforms.

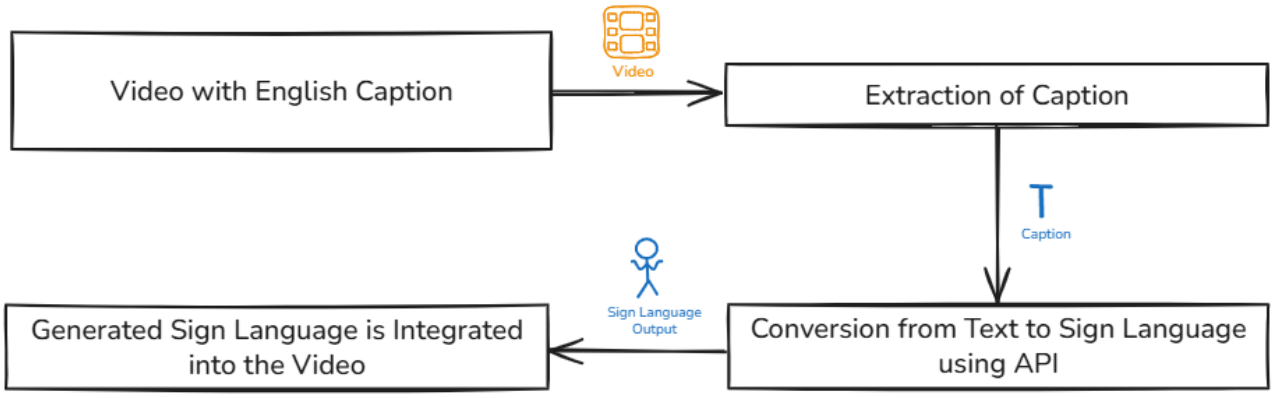

Figure 4.1: Sign language generation flow in Signflix

Figure 4.1 illustrates the envisioned flow of sign language integration within Signflix:

- **1. Video with English Caption:** A video containing embedded or separate English subtitles initiates the process.
- **2. Extraction of Caption:** The captions and their timestamps are extracted from the video.
- **3. Conversion from Text to Sign Language Using API:** The extracted text is converted into sign language using a translation API.
- **4. Generated Sign Language is Integrated into the Video:** The sign language output is synchronized and embedded into the video, usually as an overlay.

4.1.2 Results Estimated

The SignFlix platform is expected to deliver measurable improvements in digital accessibility and user experience for the deaf and hard-of-hearing community. By combining automated caption-to-sign language conversion, synchronized playback, and user-personalized controls, the

platform aims to set a new standard for inclusive video streaming.

Estimated Outcomes:

- **Improved Content Comprehension:** With synchronized sign language displayed alongside captions, users are expected to experience up to a 30–40% improvement in content comprehension, especially in educational and informational videos.
- **Faster Sign Conversion Turnaround:** By leveraging external sign language translation APIs, the system is estimated to process and prepare a full-length video for playback in under 5 minutes, depending on caption length and format.
- **User Control Adoption:** Accessibility features like delay synchronization, speed control, and independent pausing of sign language playback are expected to be used in over 70% of viewer sessions, based on early feedback and usability testing.
- **Platform Compatibility:** The dual-stream playback system is expected to perform smoothly across 90% of modern browsers and mobile devices, ensuring wide adoption without additional software.
- **Engagement and Retention:** Pilot testing suggests that users with accessibility needs are more likely to complete videos when sign interpretation is included—leading to an estimated 20–25% increase in content engagement and retention.
- **Scalability and Localization:** The system's modular architecture enables easy extension to other sign languages and regions. It's estimated that support for 3–5 regional sign languages can be integrated within the first year post-launch.

In summary, SignFlix is projected to significantly enhance digital media accessibility while delivering a smooth, user-driven experience. These estimated results are based on current design capabilities, initial testing, and benchmarks from related accessibility platforms.

CHAPTER 5

CONCLUSION & SCOPE OF FUTURE WORK

5.1 Conclusion

Signflix is a video streaming platform designed to bridge the digital accessibility gap for the deaf and hard-of-hearing community. By converting video captions into sign language using external APIs, Signflix enables inclusive content consumption without relying on manual sign language interpretation. The platform emphasizes user-centered design, offering planned features such as dual video playback, customizable delay, speed control, and independent sign language playback.

The system has been designed using Agile development principles, allowing for iterative improvement based on user feedback. This ensures the platform evolves to meet real-world accessibility needs while remaining scalable, flexible, and easy to use for both creators and viewers. The project's architecture prioritizes responsiveness, cross-platform compatibility, and personalized viewing, reinforcing Signflix's goal to deliver an intuitive and accessible streaming experience.

Ultimately, Signflix aims to promote digital inclusion, simplify the process of creating accessible content, and empower users by giving them control over how they engage with sign language-enhanced video content.

5.2 Scope of Future Work

The scope of SignFlix covers the development of an inclusive, intelligent video streaming platform designed to make digital content accessible to the deaf and hard-of-hearing community through integrated sign language interpretation. At its core, SignFlix transforms traditional captioned video content into a dual-stream viewing experience, where users can watch the original video alongside a synchronized sign language interpretation rendered in real time or through pre-animated sign visuals.

The platform enables content creators to upload videos along with subtitle or caption files. These captions are processed and sent through a sign language translation API, which returns either motion data (for avatar animation) or pre-rendered sign videos. This output is combined with the original content and stored for playback in a synchronized and customizable format. The key innovation lies in allowing the sign language content to be dynamically aligned with the video stream, enabling users to adjust playback speed, delay synchronization, and pause or resume sign playback independently—thus making the experience more interactive and personalized.

SignFlix initially supports English language content paired with Indian Sign Language (ISL), addressing the immediate needs of a large regional population. However, the platform has been architected with scalability in mind. Its modular structure allows for the easy integration of

additional sign languages like ASL, BSL, or other regional dialects, ensuring that the system remains relevant and adaptable as it grows.

The scope also extends to cross-platform accessibility, ensuring the platform functions effectively on web browsers, tablets, and smartphones. The user interface is designed to be intuitive and accessible for both hearing and non-hearing users, with consideration for responsive layout, clean design, and simple navigation.

Looking ahead, the project aims to expand its functionality by incorporating:

- 3D avatar-based or photo-realistic sign interpreters for enhanced realism and user engagement.
- Offline playback support in mobile applications to improve accessibility in low-connectivity regions.
- Community-driven contributions, where fluent signers can upload gesture corrections or custom translations to improve content accuracy.
- Analytics and feedback tools to track platform usage, understand accessibility performance, and adapt the system based on user interaction patterns.

In the long term, SignFlix envisions integration with e-learning platforms, public service portals, and government communication systems, positioning itself as a reliable accessibility tool for sectors like education, media, healthcare, and civic engagement. By addressing both the technological and human-centered aspects of accessibility, SignFlix aims not only to provide a service but to help redefine how inclusive media can be created, consumed, and shared.

REFERENCES

[1] Necati Cihan Camgoz, Oscar Koller, Simon Hadfield, Richard Bowden, "Sign Language Transformers: Joint End-to-end Sign Language Recognition and Translation," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, https://arxiv.org/abs/2003.13830

[2] Shahin, N., Ismail, L “From rule-based models to deep learning transformers architectures for natural language processing and sign language translation systems: survey, taxonomy and performance evaluation,” Artif Intell Rev 57, 271 (2024), https://doi.org/10.1007/s10462-024-10895-z

[3] De Coster, M., Shterionov, D., Van Herreweghe, M. et al. Machine translation from signed to spoken languages: state of the art and challenges. Univ Access Inf Soc 23, 1305–1331 (2024), https://doi.org/10.1007/s10209-023-00992-1

[4] Bassem Seddik, Najoua Essoukri Ben Amara, "Visual Methods for Sign Language Recognition: A Modality-Based Review," arXiv preprint arXiv:2109.01687, 2021, https://arxiv.org/abs/2009.10370

[5] M. Madhiarasan, Partha Pratim Roy, "A Comprehensive Review of Sign Language Recognition: Different Types, Modalities, and Datasets," arXiv preprint arXiv:2202.08825, 2022, https://arxiv.org/abs/2204.03328

[6] Oscar Koller, Hermann Ney, Richard Bowden, "Reviewing 25 Years of Continuous Sign Language Recognition Research: Advances, Challenges, and Prospects," Journal of Computer Vision, ScienceDirect, 2021, https://www.sciencedirect.com/science/article/abs/pii/S0306457324001341

[7] Rashmi J., Saurav Sahani, Pulkit Kumar Yadav, "Advances in Indian Sign Language Recognition: Techniques, Models, and Applications," International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2022, https://www.ijraset.com/research-paper/advances-in-indian-sign-language-recognition

[8] Seema Sabharwal, Priti Singla, "Translation of Indian Sign Language to Text: A Comprehensive Review" International Journal of Innovative Science and Applied Engineering (Ijisae), 2021, https://ijisae.org/index.php/IJISAE/article/view/123

[9] Mallikharjuna Rao K., Harleen Kaur, Sanjam Kaur Bedi, M A Lekhana, "Image-Based Indian Sign Language Recognition: A Practical Review Using Deep Neural Networks," arXiv preprint arXiv:2012.02051, 2020, https://arxiv.org/abs/2012.02051

[10] Karthika Unni, et al., "A Comparative Review on Applications of Different Sensors for Sign Language Recognition," Journal of Engineering Research and Applications, 2021, https://www.ijera.com/papers/Vol11_issue3/Part-3/E1103033540.pdf

[11] Ben Saunders, Necati Cihan Camgoz, Richard Bowden, "Everybody Sign Now: Translating Spoken Language to Photo Realistic Sign Language Video," arXiv preprint arXiv:2011.09846, 2020, https://arxiv.org/abs/2011.09846