In-depth lip interpretation Technology (1).docx

by Theresa Thomas

In-Depth Lip Interpretation Technology

Tanishq Arora
Apex Institute of Technology-Artificial Intelligence
Chandigarh University
Mohali, Punjab
21bcs11285@cuchd.in

Pragya Singh
Apex Institute of Technology-Artificial Intelligence
Chandigarh University
Mohali, Punjab
21bcs8786@cuchd.in

Sanskar Sharma
Apex Institute of Technology-Artificial Intelligence
Chandigarh University
Mohali, Punjab
21bcs6601@cuchd.in

Kavya Arora
Apex Institute of Technology-Artificial Intelligence
Chandigarh University
Mohali, Punjab
21bcs9111@cuchd.in

Abstract —Advanced Lip Interpretation Technology delves into the leading edge of interpreting verbal communication via lip motions. This summary scrutinizes the progression of lip-reading technologies, spanning from conventional approaches to state-of-the-art AI-powered systems. It delves into their diverse uses across domains like aiding communication for the hearing impaired, bolstering security protocols via voice verification, and supporting medical practitioners in identifying speech-related ailments. While progress has been made, obstacles remain, including the need for greater precision and ensuring inclusivity across varied demographics. This investigation highlights the potential societal ramifications, advocating for sustained research and innovation to fully leverage lip interpretation technology for broader societal advantages..

Keywords—Lip Interpretation, Lip-reading Technology, Spoken Language Deciphering, AI-driven Systems, Communication Assistance, Hearing Impaired, Speech Authentication, Medical Diagnosis, Speech Disorders, Accessibility, Societal Impact, Research and Development

I. Introduction

A. Problem Definition

The challenge posed by "In-depth lip interpretation technology" involves identifying and tackling hurdles associated with understanding and interpreting lip motions comprehensively. Here's a concise overview of the problem definition:

1. Limited Communication Accessibility: Many individuals, including those with speech disabilities or hearing impairments, rely on lip reading for communication. However, traditional methods often fall short in accurately capturing subtle lip movements.

2. Ambiguity and Context: Lip movements can be ambiguous and context-dependent. This ambiguity makes it difficult for individuals to accurately interpret speech solely through lip reading, as similar words and phrases may appear alike on the lips.

3. Variability Across Speakers: People exhibit diverse lip shapes, sizes, and movements, resulting in variability in lip patterns. This variability poses a significant challenge for developing accurate and reliable lip interpretation technologies.

4. Real-time Interpretation: Real-time interpretation of lip movements is crucial for effective communication in various settings. However, existing technologies may struggle to provide timely and accurate interpretations, particularly in noisy environments.

5. Integration with Assistive Technologies: Lip interpretation technology must seamlessly integrate with existing assistive technologies, such as speech-to-text systems or hearing aids. This integration aims to enhance communication accessibility for individuals with diverse needs.

Addressing these challenges necessitates advanced machine learning algorithms, computer vision techniques, and interdisciplinary research collaborations. The goal is to develop robust and reliable lip interpretation technologies capable of accurately deciphering lip movements across various contexts and environments.

B. Problem Overview

Meeting diverse communication needs, lip reading emerges as a vital alternative for individuals with speech impairments, hearing deficiencies, or amidst noisy environments where traditional methods like sign language or written communication may falter. However, its efficacy is hindered by limitations in accuracy and speed. The complexity of lip movements, involving intricate muscle motions and configurations, necessitates sophisticated analysis and an understanding of phonetic and linguistic aspects. Moreover, teaching and learning lip reading skills pose challenges due to the variability in lip patterns and the absence of standardized methods, demanding ongoing training for proficiency maintenance. Addressing these hurdles calls for technological innovations capable of providing real-time interpretation and comprehension assistance. Through advanced algorithms and interdisciplinary collaborations, such solutions aim to enhance communication accessibility and effectiveness, bridging gaps for individuals with diverse communication needs..

I. LITERATURE SURVEY

Addressing the diverse communication needs of individuals with speech impairments, hearing deficiencies, or those navigating noisy environments, lip reading emerges as a crucial alternative to traditional methods like sign language or written communication, which may prove inadequate in certain contexts. However, the efficacy of lip reading is hindered by its limitations in both accuracy and speed, posing significant challenges for effective communication. The intricate nature of lip movements, characterized by subtle muscle motions and configurations, underscores the complexity involved in deciphering them accurately, necessitating advanced analysis techniques and a comprehensive understanding of the phonetic and linguistic aspects of speech production. Additionally, the process of teaching and acquiring lip reading skills presents its own set of hurdles, including the variability in lip patterns and the absence of standardized educational approaches. This highlights the ongoing need for continuous training to maintain proficiency in lip reading.

Overcoming these challenges requires the development of innovative technological solutions that can provide real-time interpretation and comprehension assistance. Through the integration of advanced algorithms and interdisciplinary collaborations, these solutions aim to enhance communication accessibility and effectiveness, ultimately bridging gaps for individuals with diverse communication needs. By leveraging cutting-edge technologies, such as machine learning and computer vision, these solutions have the potential to revolutionize the field of lip interpretation, offering improved accuracy, speed, and reliability in deciphering lip movements. Furthermore, by providing tailored support and assistance to users in real-time, these technologies can empower individuals with speech impairments, hearing deficiencies, or other communication challenges to engage more fully and effectively in various social and professional settings.

In addition to technological advancements, addressing the challenges associated with lip reading also requires a concerted effort to raise awareness and promote understanding of the needs and experiences of individuals who rely on this form of communication. This includes advocating for greater inclusion and accessibility in educational, workplace, and public settings, as well as supporting ongoing research and development efforts aimed at further enhancing the capabilities of lip interpretation technologies. By working together to overcome these challenges, we can create a more inclusive and supportive society where individuals of all abilities can communicate and interact with confidence and dignity.

Furthermore, the integration of lip interpretation technology with existing assistive technologies, such as speech-to-text systems or hearing aids, holds promise for enhancing communication accessibility even further. By seamlessly integrating with these technologies, lip interpretation solutions can provide a comprehensive communication solution for individuals with diverse needs, ensuring that they can effectively engage in various social, professional, and educational contexts. Additionally, the development of user-friendly interfaces and customizable features can further enhance the usability and effectiveness of these technologies, allowing users to tailor their experience to meet their specific communication preferences and requirements.

Moreover, ongoing research and development efforts are essential for advancing the capabilities of lip interpretation technologies and addressing the remaining challenges in this field. This includes exploring new algorithms and techniques for analyzing and interpreting lip movements, as well as conducting studies to better understand the cognitive and linguistic processes involved in lip reading. By gaining insights into these processes, researchers can develop more accurate and reliable lip interpretation algorithms that can effectively decipher speech from lip movements in various contexts and environments.

Collaboration between researchers, engineers, healthcare professionals, educators, and individuals with lived experience is crucial for driving progress in the field of lip interpretation technology. By working together, these stakeholders can share insights, resources, and expertise to develop innovative solutions that meet the diverse needs of individuals with communication challenges. Additionally, partnerships with industry stakeholders and policymakers can help facilitate the translation of research findings into real-world applications, ensuring that lip interpretation technologies reach those who need them most.

.

II. PROPOSED SYSTEM

Creating My Dataset:

To form my dataset, I embark on the following steps:
1. Utilize my computer's webcam alongside OpenCV for streaming my speech.
2. Apply transfer learning and execute a pre-trained model (DLIB's Face Detection model) to isolate the mouth segment from the video stream.
3. Segment the video feed into frames and conduct image processing on each frame (Refer below for the image processing steps applied for our model).
4. Save the resultant video clip as a numpy array. The final dataset takes the form of a numpy array with dimensions (681, 22, 80, 112, 3), where 681 denotes the number of video clips, 22 signifies frames per video clip, and frames are of size 80x112 with 3 channels for each pixel.

Image Processing:
Image processing proves instrumental in enhancing model metrics and results. Apt processing also aids in expediting training time while curbing memory and CPU usage. Here's the breakdown of techniques employed for frame processing:

1. Lip Segmentation
2. Gaussian Blurring
3. Contrast Stretching
4. Bilateral Filtering
5. Sharpening

Lip Segmentation:

Targeting lip movements translation into words, I aim to isolate solely the lips and surrounding area. Establishing constant values for width and height, I segment out the lips and subsequently pad the segmented image to align with predefined constant dimensions.

Gaussian Blurring:
Gaussian Blurring emerges as a valuable image pre-processing technique. It involves passing a Gaussian filter through an image to blur edges and diminish contrast. This softening of sharp curves facilitates smoother calculations in the hidden layers, thereby reducing training time and overfitting risks.

Contrast Stretching:
Contrast Stretching serves as an image enhancement technique, amplifying the contrast by stretching the intensity values range. This augmentation of contrast between darker and lighter pixels enhances details' visibility, aiding the model's performance, especially in distinguishing and emphasizing lips amidst varying intensities.

Bilateral Filtering:
Bilateral Filtering, a non-linear image filtering technique, smoothens the image while preserving edges by considering both spatial and intensity distances between pixels. This noise reduction technique enhances image quality, thus bolstering model performance by eliminating unwanted noise or objects.

Sharpening:
Sharpening, an image enhancement technique, accentuates object edges, making them more defined and distinct. By augmenting edge contrast, this technique aids in object recognition, ultimately enhancing model accuracy in identifying and distinguishing features.

Deep Learning Models:
Given the classification task of video clips or frame sequences, I opt for 3D convolutional neural networks (3D CNN) to capture spatiotemporal features effectively.

3D CNN:
Upon preparing the video dataset, I craft the model architecture to capture spatiotemporal features of video frames. Commencing with three Conv3D layers using a 3x3x3 kernel size and 8, 32, and 256 filters respectively, each Conv3D layer (except the last) is succeeded by a MaxPooling3D layer with a 2x2x2 pooling size. Subsequently, a Flatten layer prepares the input for densely connected layers. Following are three Dense layers with 1024, 256, and 64 neurons respectively, interspersed with Dropout layers having a 0.5 dropout rate between each Dense layer. Finally, a Dense layer with 13 neurons concludes the architecture, representing the 13 video classification classes. The model encompasses approximately 200 million training parameters.

To mitigate overfitting, I incorporate L2 regularization into each Conv3D layer alongside multiple Dropout layers between Dense layers. L2 regularization introduces a penalty term to the loss function, reducing the impact of weights during training and promoting smaller weights to prevent over-sensitivity to training data. Dropout facilitates by randomly dropping out a fraction of neurons during training, preventing over-reliance on any single neuron and fostering robust feature learning for accurate predictions on new data.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv3d (Conv3D)             (None, 20, 78, 110, 8)    656

 max_pooling3d (MaxPooling3D  (None, 10, 39, 55, 8)    0
 )

 conv3d_1 (Conv3D)           (None, 8, 37, 53, 32)     6944

 max_pooling3d_1 (MaxPooling  (None, 4, 18, 26, 32)    0
 3D)

 conv3d_2 (Conv3D)           (None, 2, 16, 24, 256)    221440

 flatten (Flatten)           (None, 196608)            0

 dense (Dense)               (None, 1024)              201327616

 dropout (Dropout)           (None, 1024)              0

 dense_1 (Dense)             (None, 256)               262400

 dropout_1 (Dropout)         (None, 256)               0

 dense_2 (Dense)             (None, 64)                16448

 dropout_2 (Dropout)         (None, 64)                0

 dense_3 (Dense)             (None, 13)                845

=================================================================
Total params: 201,836,349
Trainable params: 201,836,349
Non-trainable params: 0
_________________________________________________________________
```

IV. METHODOLOGY

1. **Data Collection and Annotation**: A diverse dataset of lip movement videos will be compiled, spanning different speakers, languages, and contexts. These will be annotated with corresponding transcripts or phonetic information, facilitating supervised learning algorithms.

2. **Machine Learning Algorithms**: Various machine learning techniques, including deep learning models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), will be employed to train the system in recognizing and interpreting lip movements. Transfer learning will be utilized to adapt pre-trained models to lip reading tasks.

3. **Feature Extraction and Representation**: Relevant features such as shape, texture, and motion characteristics will be extracted from lip movement videos. Different feature representation methods, including handcrafted

features and learned representations, will be explored to capture the unique aspects of lip movements.

4. **Temporal Modeling**: Given that lip movements occur over time, temporal modeling techniques like Long Short-Term Memory (LSTM) networks or Temporal Convolutional Networks (TCNs) will be employed to capture temporal dependencies and dynamics in lip sequences.

5. **Integration of Context and Language Models**: Contextual information and language models will be integrated to enhance the accuracy and robustness of lip interpretation. Language models will provide additional context and constraints to improve recognition, especially for ambiguous or context-dependent lip movements.

6. **Evaluation Metrics**: Appropriate evaluation metrics such as word error rate (WER) or phoneme error rate (PER) will be defined to assess system performance. Thorough evaluations using held-out validation sets and benchmark datasets will be conducted to measure accuracy, robustness, and generalization capabilities.

7. **User-Centered Design and Evaluation**: End-users, including individuals with speech disabilities or hearing impairments, will be involved throughout the development process. Feedback will be solicited, usability studies conducted, and the system iteratively refined based on user input to ensure it meets their needs and preferences.

8. **Privacy and Ethical Considerations**: Privacy-preserving measures such as data anonymization and encryption will be implemented to protect user data confidentiality. Ethical guidelines will be adhered to, and informed consent obtained from participants involved in data collection and evaluation studies.

9. **Interdisciplinary Collaboration**: Collaboration between researchers, engineers, clinicians, linguists, and end-users will be fostered to leverage diverse expertise and perspectives in technology development. Knowledge sharing and interdisciplinary dialogue will address complex challenges and optimize system performance.

Through the adoption of these methodologies, researchers and practitioners aim to advance the development of in-depth lip interpretation technology, contributing to enhanced communication accessibility for individuals with diverse communication needs.

V. RESULT AND DISCUSSIONS

Metrics:

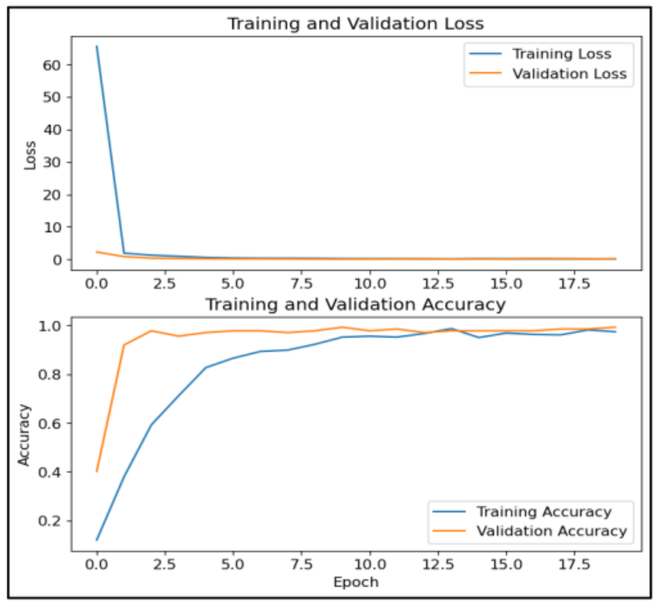
Accuracy: This fundamental metric assesses the model's performance on new, unseen data by comparing correct predictions to actual labels.

Balanced Accuracy: Similar to accuracy but considers variations in data distribution, particularly valuable for unbalanced datasets, though in this project, the dataset is mostly balanced.

Precision: Indicates the confidence level of the model by measuring the ratio of True Positives to the sum of False Positives and True Positives.

Recall: Evaluates the ratio of True Positives to the sum of False Negatives and True Positives.

F1 Score: A weighted average of precision and recall, serving as a compromise between the two metrics, with a higher score indicating a better model performance.

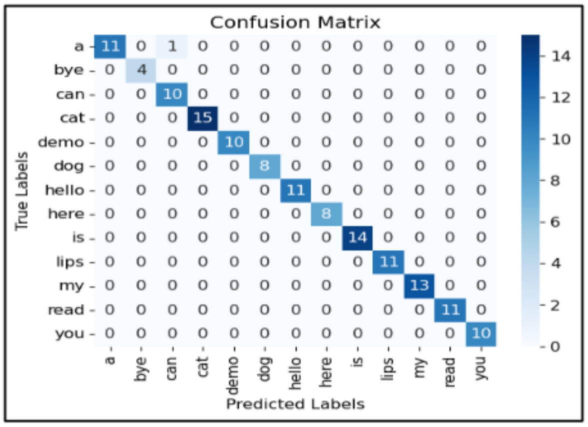


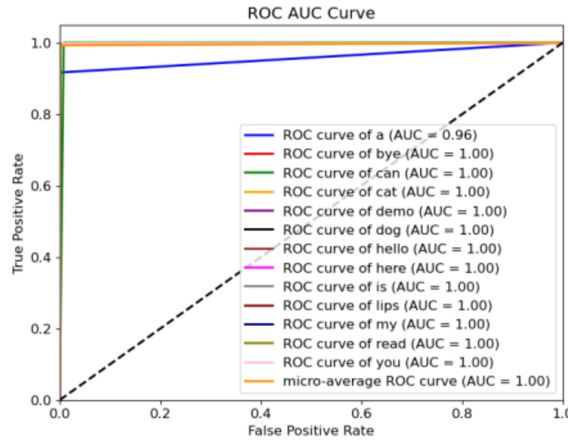
Model Training:

The 3D CNN was trained for 20 epochs, with a final training accuracy of 97.4% and testing accuracy of 99.2%, crucial for assessing performance on unseen data.

Confusion Matrix Analysis:

Examining the confusion matrix reveals instances where the model confuses visually similar words, such as "a" with "cat," indicating challenges in distinguishing certain lip movements.





Model Performance Metrics:

Precision, recall, and F1 score of 1.0 for most classes demonstrate exceptional accuracy in identifying positive samples. The balanced accuracy of 99.3% further attests to the model's robustness.

Class	Precision	Recall	F1-Score	Support
a	1.0	0.92	0.96	12
bye	1.0	1.0	1.0	4
can	0.91	1.0	0.95	10
cat	1.0	1.0	1.0	15
demo	1.0	1.0	1.0	10
dog	1.0	1.0	1.0	8
hello	1.0	1.0	1.0	11
here	1.0	1.0	1.0	8
is	1.0	1.0	1.0	14
lips	1.0	1.0	1.0	11
my	1.0	1.0	1.0	13
read	1.0	1.0	1.0	11
you	1.0	1.0	1.0	10

ROC AUC Curve:

The ROC curve illustrates consistently high AUC values, indicating strong discriminatory ability. The curve's steep ascent towards the top-left corner signifies high true positive rates at low false positive rates, indicative of the model's effectiveness.

VI. CONCLUSION

The goal of this research is to create an algorithm that translates lip motions into spoken words through the combination of computer vision and deep learning. A dataset with seven hundred video clips depicting spoken phrases is created by using transfer learning. Several image processing techniques, comprising bilateral filtering, sharpening, contrast stretching, Gaussian blurring, and lip segmentation, have been applied to each clip. Libraries like scikit-learn, TensorFlow, Keras, OpenCV, PIL, and numpy have been utilized in data preparation and model training. Convolutional neural networks and dense layers make up the model architecture that somewhat achieves noteworthy training and validation accuracies of 99.2% and 97.4%, respectively. Furthermore, the model is lightweight, making it possible for real-time word identification after training.

VII.References

[1] A. Garg, J. Noyola, S. Bagadia, "Lip reading using CNN and LSTM," in Technical Report, 2016.

[2] Y. Li, Y. Takashima, T. Takiguchi, Y. Ariki, "Lip reading using a dynamic feature of lip images and convolutional neural networks," in 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), pp. 1–6, June 2016.

[3] S. Petridis, Z. Li, M. Pantic, "End-to

-end visual speech recognition with LSTMs," CoRR, vol. abs/1701.05847, 2017.

[4] A. Rekik, A. Ben-Hamadou, W. Mahdi, "A new visual speech recognition approach for RGB-D cameras," in Image Analysis and Recognition (A. Campilho and M. Kamel, eds.), (Cham), pp. 21–28, Springer International Publishing, 2014

[5]Takeshi Saitoh, Ryosuke Konishi ―A Study of Influence of Word Lip Reading by Change of Frame Rate‖. ISCA Archive on Audio-Visual Speech Processing :speech.org/ archiveHakone, Kanagawa, Japan"

[6]Takeshi Saitoh ―Development of Communication Support System Using Lip Reading‖ IEEJ Transactions On Electrical And Electronic Engineering, IEEJ Trans 2013; 8: 574–579 Published online in Wiley Online Library (wileyonlinelibrary.com). DOI:10.1002/tee.21898

[7] Cruz, Hans Miguel Puente, Jofet Kane T Santos, Christian, Vea Larry A., Rajendaran Vairavan ―Lip Reading Analysis of English Letters as Pronounced by Filipino Speakers Using Image Analysis‖ 1st International Conference on Green and Sustainable Computing (ICoGeS) 2017 IOP PublishingIOP Conf. Series: Journal of Physics: Conf. Series 12345678901019 (2018) 012041 doi :10.1088/1742-6596/1019/1/01204

[8] I. Anina, Z. Zhou, G. Zhao, M. Pietikainen, "Ouluvs2: A multi-view audiovisual database for non-rigid mouth motion analysis," in 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), vol. 1, pp. 1–5, May 2015.

[9] E. K. Patterson, S. Gurbuz, Z. Tufekci, J. N. Gowdy, "Moving talker, speaker-independent feature study, and baseline results using the CUAVE multimodal speech corpus," EURASIP J. Appl. Signal Process., vol. 2002, pp. 1189–1201, Jan. 2002.

[10] W. Dong, R. He, S. Zhang, "Digital recognition from lip texture analysis," in 2016 IEEE International Conference on Digital Signal Processing (DSP), pp. 477–481, Oct 2016.

[11] T. Stafylakis, G. Tzimiropoulos, "Combining residual networks with LSTMs for lipreading," CoRR, vol. abs/1703.04105, 2017.

[12] J. S. Chung, A. Zisserman, "Lip reading in the wild," in Asian Conference on Computer Vision, pp. 87–103, Springer, 2016.

[13] Y. Takashima, R. Aihara, T. Takiguchi, Y. Ariki, N. Mitani, K. Omori, K. Nakazono, "Audio-visual speech recognition using bimodal trained bottleneck features for a person with severe hearing loss," in INTERSPEECH, 2016.

[14]Sunil S. Morade Suprava Patnaik (2015), ―Comparison of classifiers for lip reading with CUAVE and TULIPS database‖. Optik - International Journal for Light and Electron Optics, 126(24), 5753–5761.

[15] M.H. Rahmani, F. Almasganj Lip-reading via a DNN-HMM hybrid system using combination of the image-based and model-based features

[16] T. Stafylakis, G. Tzimiropoulos, Combining Residual Networks with LSTMs for Lipreading, in: Proc Annu Conf Int Speech Commun Assoc INTERSPEECH 2017-August, 2017, 3652–3656.

[17] G. Sterpu, H. Naomi, Towards Lipreading Sentences with Active Appearance Models. AVSP, 2017, 70–75.

[18] K. Thangthai, R. Harvey, Improving computer lipreading via DNN sequence discriminative training techniques, in: Proc Annu Conf Int Speech Commun Assoc INTERSPEECH 2017-August, 2017, pp. 3657–3661. https://doi.org/10.21437/INTERSPEECH.2017-106.

[19] K. Xu, D. Li, N. Cassimatis, X. Wang LCANet: End-to-end lipreading with cascaded attention-CTC

[20] M. Wand, J. Schmidhuber, N.T. Vu, Investigations on End- to-End Audiovisual Fusion. ICASSP, IEEE Int Conf Acoust Speech Signal Process – Proc 2018-April

[21] S. Petridis, T. Stafylakis, P. Ma, et al., End-to-End Audiovisual Speech Recognition. ICASSP, IEEE Int Conf Acoust Speech Signal Process – Proc 2018-April, 2018

[22] T. Afouras, J.S. Chung, A. Zisserman, Deep Lip Reading: a comparison of models and an online application, in: Proc Annu Conf Int Speech Commun Assoc INTERSPEECH 2018-September, 2018, pp. 3514–3518.

[23] S. Petridis, Y. Wang, Z. Li, M. PanticEnd-to-End Audiovisual Fusion with LSTMsThe 14th International Conference on Auditory-Visual Speech Processing. International Speech Communication Association (2018), pp. 36-40

[24] T. Saitoh, Z. Zhou, G. Zhao, M. Pietikäinen, Concatenated Frame Image Based CNN for Visual Speech Recognition. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 10117 LNCS, 2016, pp. 277–289. https://doi.org/10.1007/978-3-319-54427-4_21.

[25] J.S. Chung, A. Zisserman, Lip Reading in the Wild. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 10112 LNCS, 2016, 87–103. https://doi.org/10.1007/978-3-319-54184-6_6.

[26] K. Thangthai, R.W. Harvey, S.J. Cox, B.J. Theobald, Improving lip-reading performance for robust audiovisual speech recognition using DNNs, in: AVSP, 2015, pp. 127–131

[27] M. Zimmermann, M. Mehdipour Ghazi, H.K. Ekenel, J.-P. Thiran, Visual Speech Recognition Using PCA Networks and LSTMs in a Tandem GMM-HMM System. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 10117 LNCS, 2016, 264–276. https://doi.org/10.1007/978-3-319-54427-4_20.

[28] Lee D, Lee J, Kim K-E (2016) Multi-view Automatic Lip-Reading Using Neural Network. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 10117 LNCS:290–302. https://doi.org/10.1007/978-3-319-54427-4_22.

[29] Koumparoulis A, Potamianos G (2019) Deep View2View Mapping for View-Invariant Lipreading. 2018 IEEE Spok Lang Technol Work SLT 2018 - Proc 588–594. https://doi.org/10.1109/SLT.2018.8639698.

[30] Han H, Kang S, Yoo CD (2018) Multi-view visual speech recognition based on multi task learning. Proc - Int Conf Image Process ICIP 2017-September:3983–3987. https://doi.org/10.1109/ICIP.2017.8297030.

[31] A. Bakry, A. Elgammal MKPLS: Manifold kernel partial least squares for lipreading and speaker identification Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit, 684–691 (2013)

In-depth lip interpretation Technology (1).docx

ORIGINALITY REPORT

7%	4%	3%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Manipal University Student Paper	1%
2	Submitted to Heriot-Watt University Student Paper	1%
3	Submitted to California State University, Sacramento Student Paper	1%
4	Submitted to PES University Student Paper	1%
5	ijsrst.com Internet Source	1%
6	ai-sdc.github.io Internet Source	<1%
7	G. Aswanth Kumar, Jino Hans William. "Development of Visual-Only Speech Recognition System for Mute People", Circuits, Systems, and Signal Processing, 2021 Publication	<1%
8	dyuthi.cusat.ac.in Internet Source	<1%

9	www2.mdpi.com Internet Source	<1 %
10	Joseph Migga Kizza. "Chapter 17 Artificial Intelligence and the Security of Computing Systems", Springer Science and Business Media LLC, 2024 Publication	<1 %
11	Konda Reddy Mopuri, Aditya Ganeshan, R. Venkatesh Babu. "Generalizable Data-Free Objective for Crafting Universal Adversarial Perturbations", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019 Publication	<1 %
12	link.springer.com Internet Source	<1 %
13	"Pattern Recognition", Springer Science and Business Media LLC, 2020 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		