International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 05 May 2024, pp: 966-971 www.ijaem.net ISSN: 2395-5252

Multi-Modal Sign Language Detection: Integrating Sign, Text, Image, and Voice

Arushi Gupta, Shitiz Rajvanshi, Taniya Singh, Shubham Goel
KIET Group of Institutions
Ghaziabad, Uttar Pradesh

Date of Submission: 18-05-2024 Date of Acceptance: 28-05-2024

I. INTRODUCTION

The capabilities of Android apps have beengreatly improved, enabling Java programs torunonmobiledevices.Thankstothisadvancement, people all over theworldcannowusetheirmobiledevicestoreadandwriteemails,browsewebsites,andplayJavagames.
Taking note of this development, wesuggestingAndroidapplicationstoimprovecommunication.TheintroductionofSMSandMMSmadeiteasierfordeafpeople,whohadnot oftenusedcellphones,tocommunicateremotely. Deaf people can now communicatewithbothhearinganddeafpeopleviatexts.
Even though there are deaf or muted people allaround us, many people find it difficult tocommunicate with them. There is a needfor asolution that makes communication easier foreveryonebecauseavoidinginteractionisnotasolution. We have created an application toaddressthisdemandandfacilitateusers'everydaycommunication.Evenastechnologydevelopsfurther,itsapplicationshouldconstantly aim for advantages. Our programtries to make it easier for those who are dumbor deaf to communicate with others aroundthem. While other developers have tried toimprove sign language apps, our goal is tomake ours more dependable and effective. Thefocusofcurrentsignlanguageappsistypicallyon text-to-sign or sign-to-text conversion. Onthe other hand, our program consists of twomodules:SigntoTextandTexttoSign.Furthermore, our application enables users toupload their own images, cropping them aftertakingapictureor choosingonefromthegallery.Next,theimage'stextisshownonthescreen, and a sign language version of it isproduced.audiotoSignConversion,whichconvertsaudiomemosortalkstotextandthenback to sign language, is another noteworthyfeature.Becauseiteliminates theneed fortyping, this feature is very helpful for Englishspeakers who need to translate text into signlanguage for greater understanding. Amongthe numerous difficulties confronted by thedeafanddumbisobjectrecognition.Withthehelpof oursoftware,usersmayidentifyobjects in an image by taking a picture orchoosingonefromthegallery,withouthavingtodependonothersto identifythem.Theapp alsoshowsthepercentagebywhichtheimagematches the object. Language recognitionisourapp'slastfunctionality. When usersentertextinalanguagetheyarenotacquainted with, the app will translate it intoAmericanEnglish,recognizeit,andcreatetheappropriategestureimage.

II. LITERATURE REVIEW

In[1]SignLanguageespeciallyIndianSignLanguage (ISL), for the deaf and mute. It noteslimitedresearchpost-ISLstandardization,focusingonstatichandgestureswithminimalattentiontodynamics.DespiteeffortsonISL alphabet recognition, the process involves multiplestages,surveyedtoassessresearchprogress.In [2]an intelligent system for translating sign languagetotext,comprisinghardwareandsoftwarecomponents.Thehardwareincorporatesflex,contact, and inertial sensors on a glove. Softwarefeatures a classification algorithm leveraging k-nearestneighbors,decisiontrees,anddynamictimewarping,enablingstaticanddynamicgesturerecognition.In[3]threemethodsforsub-unit-based sign recognition. Boosting is employed tolearn appearance-based sub-units, merged with asecond-stageclassifierforword-levelsignlearning.Anotherapproachintegrates2Dtracking-basedsub-unitswithappearance-basedhandshape classifiers. The final method translatestheseinto3D,enablingreal-time,user-independent recognition of isolated signs. In [4] adeepconvolutionalneuralnetworkfordirectclassificationofhandgesturesinimages,eliminatingtheneedforsegmentationordetectionstages. In [5] two novel

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 05 May 2024, pp: 966-971 www.ijaem.net ISSN: 2395-5252

hand gesture recognitionapproachesforreal-timesignlanguagecomprehension.EmployingahybridfeaturedescriptormergingSURFandHuMomentInvariantmethodsyieldsastrongrecognitionrate.SURFandmomentinvariantfeaturesexhibitresilience to diverse variations, ensuring effectivereal-timeperformance.[5]Introducestwoinnovative methods for real-time recognition ofhand gestures in sign language. These methodsmerge SURF and Hu Moment Invariant techniquesintoacombinedfeaturedescriptor,improvingrecognition accuracy while maintaining low timecomplexity. They also introduce derived featuresandutilizeKNN,SVM,andHMMforclassification,demonstrating enhanced real-timeefficiencyandrobustness.[6]Presentsnovelstrategiesforreal-timerecognition,translation,andvideoproductioninSignLanguage(SL).EmployingMediaPipeandhybridCNN+Bi-LSTM models for recognition, and NMT + GANmodelsforvideogeneration,achievingclassificationaccuracyexceeding95%.Evaluationmetricsrevealsubstantialenhancements,includinga38.06BLEUscoreandimpressivevisualquality.[7]AddressesthechallengesofContinuous Sign Language Recognition (CSLR)byintroducingSignBERT,adeeplearningframeworkmergingBERTandResNet.Outperforming conventional methods in accuracyandworderrorrateondemandingdatasets,SignBERTunderscoresitseffectivenessinmodelingsignlanguagesandextractingspatialfeatures for real-time CSLR. [8] Examines signlanguage research, particularly vision-based handgesture recognition systems from 2014 to 2020.Through analysis of 96 articles, it identifies keyresearch areas: data acquisition, environment, andgesturerepresentation.Signer-dependentrecognitionaverages88.8%,whilesigner-independentrecognitionaverages78.2%,indicatingopportunitiesforimprovement,especially in continuous gesture recognition. [9]Introducesadynamichandgesturerecognitionsystemleveragingmultipledeeplearningarchitectures. Evaluated on a challenging dataset,itoutperformsexistingmethods,demonstratingeffectiveness in uncontrolled environments withdiverse gestures. [10] Presents a real-time handgesturerecognitionsystemutilizingacost-effectivewebcamandimageprocessingtechniques.Thesystemcomprisesfourstages:imagepreprocessing,regionextraction,featureextraction,andmatching,achievinga90.19%recognitionrateforAmericanSignLanguage(ASL)alphabetgesturesundervariouslightingand handconditions.

III. SYSTEM REQUIREMENTS

A) HARDWARE REQUIREMENTS

Specifichardwarecomponentsforthemobile application are necessary for the system's overallsuccessfuloperation.Tosupportthesystem'svariousfunctions,anAndroidsmartphonerunningAndroid version 5.0 or higher is required. Thegadget ought to possess a rear camera that canrecognizemotionsinAmericanSignLanguage(ASL),enablingthevisualinputnecessaryforefficientcommunication.A microphoneisalsonecessary todetecthumanvoice.

B) SOFTWARE REQUIREMENTS

1. **AndroidStudio:**Theofficialintegrateddevelopment environment (IDE) for the Androidoperating system from Google is called AndroidStudio.SpecificallydesignedforAndroidprogramming,thisIDEisbuiltonJetBrains' IntelliJIDEAsoftware.

2. **GoogleVisionAPI:**ThisAPIisacomponentofthe Google API family and offers applicationprogramminginterfaces(APIs)forintegratingwithotherservicesandcommunicatingwitharange of Google services. Google Maps, Gmail,Search, and Translate are a few examples. TheseAPIs can be used by third-party apps to expand orimprovethefeaturesofalready-availableservices.Specifically,theGoogleVision APIprovidesfeatures including analytics, user data access, andmachinelearningasaservice(thePredictionAPI).
3. **FirebaseMLKit:**Thismobilesoftwaredevelopmentkit(SDK)makesuseofGoogle'smachinelearningknow-howtoimproveapplicationsforiOSandAndroiddevices.Fordevelopers with varying levels of experience, itprovidesarobustandintuitivepackageforintegratingmachinelearningfeatures.Itonlytakesa fewlinesofcodetoaddmachinelearningcapabilities to an app with ML Kit; developers nolonger need to be experts in neural networks ormodeloptimization
4. **HashMap Class:** The HashMap functionmakes it easier to map processed input to thedatabase that is stored. The Map interface isimplemented by the HashMap class, whichenables the storing of key-value pairs wherethe keys need to be distinct. This class, whichis part of the java.util package, is

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 05 May 2024, pp: 966-971 www.ijaem.net ISSN: 2395-5252

essential toeffectively maintaining associations betweenprocessed inputs and the relevant data kept inthedatabaseofthesystem.

IV. FLOW CHARTS AND DIAGRAMS

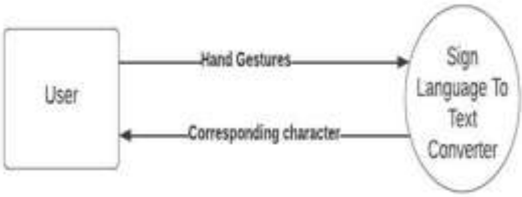

Figure1:DFD Level-0 Diagram

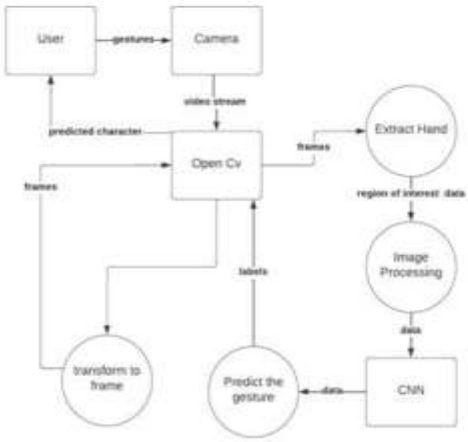

Figure2:DFDLevel-1Diagram

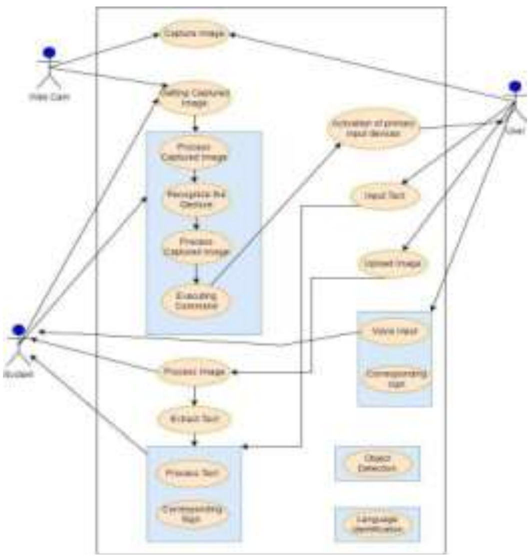

Figure3:UseCaseDiagram

V. PROPOSED SYSTEM

Forthosewhoaredeaformute,thisappoffersaground-breakingalternativethatmakescommunication easier. in the context of theirenvironment.voice-to-signtransmission,vehiclevoicerecognition,andautomatictranslation are all supported by the proposedtechnology. In contrast tothe current systemsthatmostlyconcentrateonone-waycommunication,ourproposalofferstheabilitytocommunicateinbothdirections,from sign language to text or the other wayaround.Userscan interprettextsintosignlanguage and texts into signlanguage, whichpromotesbettercommunicationandeducational opportunities.WewilluseSupportVectorMachines(SVM)forbothregression and classification techniques in ourresearch.Thismethodwillbeusedtoclassifysignsto determinehowwelltheyarefunctioning.A rich image dataset is essentialfor the implementation of sign language. Atleast20distinctrepresentationsofeachalphabetshouldbeused,consideringchangesin background, lighting(bulb, lamp, cameraflash, daylight, and dim light), and distance.Ourattentionturnstoimprovingimagerecognition for effective alphabet identificationafterthedatasetisestablished.
Text recognition will improve if conditionsfor text-to-image matching are included. Usingatext-to-signlanguageinterpreter,thetextcan be recognized andthen shown as an image oranimationinsignlanguage.Whetheranimage is taken with the camera or importedfrom the gallery, every elementin the picturewillbeexaminedinpercentagetermstoprovide specific details. We will include asectionthat translates spoken languages frommany nations,supporting as many languagesas we can. This feature facilitates seamlesscommunicationsbyenablinguserstocomprehend the language spoken by othersandensuringworldwideaccessibility.Forthoroughcommunication,eachlanguagetypewillbetranslatedintoacoherentEnglishsentenceandthenintosign language.

VI. SYSTEM DESIGN

Framework Plan is primarilypartitionedintosixsegments:

1. Textto SignChange
2. PicturetoSignTransformation
3. VoicetoSignChange
4. SigntoTextTransformation
5. ObjectDetection
6. LanguageIdentification

Thesesegmentsperformtheirexercisesutilizing threeimportant modules:

A. ContentAcknowledgmentFramework

B. FirebaseMachineLearningUnit

C. MotionAcknowledgmentFramework

TheFrameworkPlanmaybeacomprehensive system that envelops six unmistakable segments,eachservingasaspecializedaspectofthesystem'scapabilities. These segments incorporate Text toSignChange,PicturetoSignConversion,VoicetoSign Transformation, Sign to Text Change, ObjectIdentification,andLanguageIdentification.Together,theyframeacohesivedesignthataddressesdifferentmodesofinteractionandacknowledgment.

Central to this plan are three basic modules thatcollaboratetoguaranteeexactandseamlesschanges.TheContentAcknowledgmentFramework,theprimarymodule,servesasthespine of the system's capacity to interpret contentinput into expressive sign dialect signals. FirebaseMachineLearningUnit,themomentcentercomponent, essentially improves acknowledgmentcapabilitiesoverarunofinputmodes, improving thesystem'sprecisionandunwaveringquality.TheMotionAcknowledgmentFramework,thethirdmodule,playsaessentialpartintranslatingperplexinghanddevelopmentsandmotions,contributingtobothcontentandsigndialecttransformations.

TexttoSignTransformation,thesystem'sintroductory area,enablestheinterpretationofwritten literary input into meaningful visual signdialectexpressions.This handleguaranteesthatclients can easily pass on their messages in a waythat'sallaroundcaughton,bridgingthegapbetween written dialect and the nuanced dialect ofsigns. Picture to Sign Change takes after, utilizingprogressedalgorithmstohandlepicturesandcreatecomparing sign dialect motions. This include isespeciallyusefulwhenmanagingwithvisualsubstance,enhancingcommunicationthroughvisualprompts. Voice to Sign Transformation speaks to anotherfeatureoftheframework,leveragingsoundinformation to change over talked language intosigndialectrepresentations.Thiscapabilityimproves openness for people with sound-relatedimpedances,empoweringthemtolockininconversationsandpassontheircontemplationsutilizingsigndialectmotions.SigntoTextTransformationservesastheconversehandle,translating hand signals captured through camerasinto printed yield. This highlight finds applicationin circumstances where clients may lean toward tocommunicate using motions instead of written ortalked dialect.

Thesystem'scapabilitiesexpandpastphoneticcommunication, including Object detection. Thissegmentutilizescutting-edgecalculationstodistinguishandclassifyobjectsinsidepictures,upgrading the system's utility in recognizing andconnection with the encompassing environment.Theultimatesegment,LanguageIdentificationdiscourages the dialect utilized within the giveninput,showcasingthesystem'sversatilitytodifferentphoneticsettings.Consolidatingthesecapabilitiesintoanall-encompassingoutline-workis accomplished through fastidious integration ofmodules.Themeasuredplanguaranteesadaptability , versatility,andeffectiveness.Byleveraging the qualities of Content Recognition,thecapabilitiesof FirebaseMLPack,andtheaccuracyofSignalAcknowledgment,theframework conveys an upgraded client encounterthat consistently bridges abunch of input strategieswith their comparingyieldrepresentations.

TheFrameworkPlanencapsulatesamodernintegrationofsixspecializedareas,eachcontributingtoaenergeticandcomprehensiveuser experience.ThroughthesynergyofContentAcknowledgment, Firebase ML Pack, and MotionAcknowledgment,thisplanexhibitsourcommitment to saddling innovation for exact andflexible dialect and question acknowledgment. Byaddressing different modes of interaction, we pointtoformancomprehensivestagethatengagesclientsto communicate viably and exceptionless,notwithstanding of their favored communicationmodeorphoneticfoundation.

VII. IMPLEMENTATION

1. TexttoSignConversion

The text recognition system created in AndroidStudio blends classic text input with revolutionaryhand gesture recognition, providing users with aninteresting and dynamic experience. Users entertextusingfamiliartechniques,whicharesubsequentlytranslatedintoequivalenthandmotions.Imagerecognitiontechniquesandmachinelearningalgorithmsenablepreciesegesture recognition. The system converts motionsbacktowords,allowinguserstoevaluateandchangetheirinput.Benefitsincludedynamicengagement ,individualizedgestures,anddemonstrating the combination of classic and newinteractionapproaches.Overall,the systemisauniquecombinationoftraditionalinputwithcutting-edgetechnologythatimprovesuserexperienceandemphasizesthepossibilityforintuitive interactionswithdigital devices.

2. PicturetoSignConversion

The built text recognition system links the

GoogleVisionAPIwithAndroidStudio,allowingforte xtextractionfromphotosaswellasnovelhandgesture identification. The procedure begins withtheextractionoftextfromphotographsusingOCRt echnology.Eachcharacteristhenassignedacorrespon dinghandmotion,improvinguserinteraction.TheAndr oidsoftwaretakeshandmotions in real time, identifies them using imagerecognition techniques, and then maps them backto characters using machine learning algorithms.Thisintegratedsystemprovidesuserswith convenience,efficiency,andengagementwhiledemon stratingthecombinationofmoderntechnologyandcrea tiveinterfaces.Overall,itmarks a big step forward in user interface andtechnologicalintegration,simplifyingtextextracti onandimprovingtheuserexperiencethrough gesture-basedinputmethods.

3. VoicetoSignConversion

The created text recognition system uses Google'sVoiceAPIandAndroidStudiototranscribesp okenwordsintotextandincorporateshand motion recognition for interaction. The procedurebegins by using the Voice API to translate spokenwords into text, which acts as the foundation forfurther interactions. Each character is paired witha hand motion, which improves user involvement.The Android software takes spoken words in realtime,recognizesthemwithspeechrecognitiontech nology,andmapsthemtomotionsusingmachinelearni ngtechniques.Thisintegratedsystemprovidesuserswi thconvenience,efficiency, and engagement while demonstratingthecombinationofspeechandgesturere cognitioncapabilities and innovative interfaces. Overall, itmarks a huge step forward in human interactionand technology convergence, simplifying spokenword conversion and increasing user engagementthrough gesture-based inputmodalities.

4. SignToTextConversion

Theproject'sgoalistocreateaSign-to-TextConversionsysteminAndroidStudiothatcombin esreal-timehandmotionstakenbythedevice'scamerawithAm ericanSignLanguageletters.Thesysteminterpretsmo vements usingpicturerecognitionandmachinelearningtechniq ues, which are similar to computer visionprinciples. The system learns to detect individuallettersandconvertsgesturesintotextreprese ntations on the screen after being trainedwithanASL-gesturedataset.Thismethodimprovesaccessibilityan dcommunicationforsignlanguageusers,fosteringmor einclusiveinteractions.Furthermore,itdemonstratest heconvergence of contemporary picture recognitionwithAndroidappdevelopment,whichfost ersintuitive communication methods and influenceshow people interact with technology. Overall, theprojectemphasizestherelationshipbetweenpicture recognition,appdevelopment,andaccessibility,allow inguserstoefficientlycommunicateusingASLmotion sandwrittentext.

5. ObjectDetection

The project aims to create an Object Detectionsystem using Android Studio, allowing users torecord or pick photographs and reliably identifythingswithinthemusingGoogleFirebaseML Kit.The technology provides flexibility by allowinguserstoselectbetweenlivecamerafeedsandg allery photographs. Theintegration ofFirebaseMLKitprovidesrobustobjectdetectionusin gadvanced image recognition algorithms trained onvaried datasets. Detected objects are expressed asa percentage,showingthesystem'slevelofconfidence in their recognition. This percentage-based representation provides useful insights intoimagecontent,improvesuserexperience,andhas practical uses in a variety of settings. The projectdemonstrates the seamless integration of powerfulalgorithmswithuser-friendlyinterfaces,highlighting technology's ability to ease complexprocesses.

6. LanguageIdentification

The project focuses on Language Detection withGoogle Firebase ML Kit, allowing users to entertext in any language and reliably identify it withthe press of a button. The application's text inputinterface is easy and flexible to a wide range oflinguistic preferences. The integration of FirebaseML Kit allows for the examination of entered texttodetectits language,using powerfulmachinelearning algorithms trained on a sample of 144languages.Thisallowsreal-timelanguagedetection,whichimprovesaccessibilityf orlanguagelearners,travelers,andothersdealingwith multilingual information.

VIII. RESULTSANDDISCUSSIONS

TheresultsofourSignlanguagerecognitionm odelsareshowninTable1.Ourapproachinvolvestraini ngaCNNmodelforconvertingsignlanguage to text. For text-to-sign conversion, weutilize the HashMap class to map characters

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 05 May 2024, pp: 966-971 www.ijaem.net ISSN: 2395-5252

totheirrespectivesigns.Voice-to-signconversionrelies on the Google Voice API. In the case ofimage-to-sign conversion, we initially extract textfrom the image using the Google Cloud VisionAPI,followedbyconvertingthetexttothecorresponding sign.

Table1: FeatureSpecificAccuracy

Feature	Implementation	Accuracy
Sign to Text	CNN (Convolutional Neural Network)	95%
Voice to Sign	Google Voice API	91.34%
Image to Sign	Google Cloud Vision API	92.45%

IX. CONCLUSION

The study covers potential enhancements to handgesturerecognitionsystems,includinggeneralizing the system to include more gesturesand actions, as well as training the system on datafromseveraluserstoaccountforvariancesingestureexecution.Usertestingisusefulforidentifying errors in recognition accuracy, It alsodiscussesthekeytechniques,applications,andchallenges of hand gesture recognition, includinggestureacquisitionmethods,featureextraction,classification, and applications in sign languageandrobotics.Environmentalissuesanddatasetavailability are addressed, emphasizing the needfor additional research in the topic. While currentmethodshavedemonstratedgreatperformance,thereisstillopportunityforexplorationand growthofhandgesturedetectionintoothertechnical domains such as tablets, smartphones,and game consoles. Hand gesture recognition hasthepotentialtoimprovehuman-computerinteractionsbymakingthemmorenaturalandpleasurable.Thestudyalsointroducesanautomatichand-signlanguagetranslatorformute/deafpeopleanddiscussessystemrequirements and performance objectives. It goesintodetailintosoftware issuessuchassystemstartupandrecognitionalgorithms,aswellaschallengesinidentifyingambiguousmeasurementsandrecommendingtechnicalsolutions.

REFERENCES

[1]. Research paper,” A Survey on Hand GestureRecognition for Indian Sign Language”, IRJET,2016.
[2]. Researchpaper,”GestureRecognitionandMachineLearningAppliedtoSignLanguageTranslation”, Springer, 2016.
[3]. ResearchPaper”TowardsInterpretingRobotic System for Fingerspelling Recognition inRealTime”, ACM,2018.
[4]. ResearchPaper"HandGestureMovementRecognitionSystemUsingConvolutionNeural NetworkAlgorithm",InternationalResearch Journal of Computer Science (IRJCS) Issue 04,Volume6(April2019)ISSN: 2393-9842.
[5]. ”HandGestureRecognitionforSignLanguage: A New Hybrid Approach”, Researchgate, January2020
[6]. “DevelopmentofanEndtoEndDeepLearningFrameworkforSignLanguageRecognition,Translation,andVideoGeneration”,IEEEAccess,2022
[7]. “SignBERT: A BERT-Based Deep LearningFrameworkforContinuousSignLanguageRecognition”, IEEE ccess, Volume 9, JournalArticle,2021
[8]. “A Review of the Hand Gesture RecognitionSystem:CurrentProgressandFutureDirections”.IEEEAccess,Volume9,JournalArticle,2021
[9]. “DeepLearningBasedApproachforSignLanguageGestureRecognitionWithEfficientHandGestureRepresentation”,IEEEAccess,Volume8,JournalArticle,2020
[10]. ” Real-Time Static Hand Gesture Recognitionfor American Sign Language (ASL) in ComplexBackground”, SCRIP, 2012