

ENHANCING GUN DETECTION BY USING YAMNET AUDIO CLASSIFICATION

MS.M.ANITHA¹, MS.K.BABY RAMYA², K. BINDU SAMYUKTHA³

¹ HOD & Assistant professor, Department of Master of Computer Applications, SRK Institute of Technology, Vijayawada, Andhra Pradesh

² Assistant professor, Department of Master of Computer Applications, SRK Institute of Technology, Vijayawada, Andhra Pradesh

³ MCA Student, Department of Master of Computer Applications, SRK Institute of Technology, Vijayawada, Andhra Pradesh

ABSTRACT

Gun detection in audio signals has been a critical field of study for increasing public safety and security systems. Traditionally, gunshot detection systems focused on simple acoustic characteristics and rule-based approaches, which often failed to distinguish gunshots from other loud sounds in complicated settings. These old approaches, although successful in controlled conditions, have substantial limits in terms of accuracy and scalability, particularly in loud metropolitan contexts where separating gunshots from background noises is difficult. As security concerns grow, the need for more dependable and precise gun detection systems becomes critical. The issue is the capacity to identify gunshots in real time, under varied environmental circumstances, and with a low percentage of false positive. To overcome these issues, advances in deep learning and audio categorisation have given intriguing answers. The relevance of this study stems from its potential to improve gunshot detection skills, resulting in faster and more accurate reactions in crucial circumstances. The suggested system seeks to identify gunshot noises using sophisticated audio classification algorithms, which will improve detection accuracy and resilience. Using cutting-edge algorithms trained on massive audio datasets, the system can distinguish between gunshots and other similar noises, even in loud or complicated acoustic settings. This strategy has the potential to drastically minimise false alarms while enhancing the overall performance of real-time gunshot detection systems, resulting in safer communities and quicker emergency response times, and enhances healthcare accessibility, reducing mortality rates linked to RBC infections.

Keywords: Public safety, Security, Gunshot Detection skills.

1. INTRODUCTION

1.1 Background and Overview

Gunshot detection systems have advanced dramatically over time, spurred by the growing need for enhanced public safety and faster emergency responses. Gun violence is a prominent problem in India, with reports of firearm-related crimes and terrorist acts dominating headlines. According to recent statistics, there has been an increase in gun-related crimes in metropolitan areas, highlighting the need for more dependable security measures. Traditional gunshot detection systems depended on physical sensors such as microphones or acoustic sensors installed in specified areas, but their accuracy and coverage were limited. To address these issues, machine learning and deep learning methods have been added into gunshot detection throughout time. These developments have

significantly increased detection capabilities, making systems more adaptive and scalable. Gunshot detection systems have a wide variety of applications, from public safety monitoring in metropolitan areas to incorporation with law enforcement and emergency response systems, allowing for quicker identification of gunshot occurrences and more accurate responses.

1.2 Problem Definition.

Prior to the emergence of machine learning, gunshot detection systems had various obstacles. Traditional systems depended heavily on acoustic sensor networks, which were often restricted in their capacity to detect gunshots from other loud sounds. These technologies generated a large number of false positives, particularly in loud metropolitan contexts. Furthermore, the complexity of auditory waves in various situations made real-time detection difficult. Systems were not adaptable enough to deal with variable noise levels, distances, or weapon kinds, making them less trustworthy in dynamic environments. Furthermore, conventional systems lacked scalability and could only cover a restricted region, making them unsuitable for large-scale or distant applications.

1.3 Research Motivation

The fundamental reason for this study is the rising need for precise, real-time gunshot detection technologies that may dramatically enhance public safety and emergency response times. Traditional techniques have been ineffective owing to their dependence on simple, rule-based approaches and poor environmental flexibility. With an increase in gun violence and associated occurrences worldwide, the need for more effective remedies has become critical. This study aims to improve the accuracy of gunshot detection in noisy and complex environments by incorporating advanced machine learning techniques, specifically deep learning-based audio classification, allowing for the rapid and effective identification of gun-related incidents and the initiation of appropriate responses.

1.4 Significance.

This initiative has tremendous importance in tackling the rising concerns of public safety and security. Gunshot detection systems serve an important role in rapidly informing authorities and law enforcement officials, minimising reaction time during violent occurrences. The use of machine learning-based audio categorisation improves the system's adaptability and reliability in a variety of situations, allowing gunshots to be recognised even in the presence of background noise. The system's enhanced accuracy will eliminate false alerts, ensuring that emergency resources are focused only on true shooting events. Furthermore, real-time identification will enable prompt action, perhaps saving lives and reducing future violence. In highly packed metropolitan areas or vast public places, the ability to identify gunshots and analyse the situation accurately is critical for preserving public order and safety. This project seeks to modernise and enhance current systems, making them more efficient and scalable for wider use.

1.5 Need

The growing incidence of gun violence in public places and metropolitan regions, which presents substantial dangers to public safety, necessitates the immediate implementation of this initiative. Traditional detection systems are not designed to react quickly in areas with background noise, limiting their usefulness. This system's real-time nature enables law enforcement and emergency responders to get quick notifications, significantly improving reaction times and boosting the possibility of averting more injury. In circumstances when every second counts, such as active shooter occurrences, the system's ability to reliably detect and pinpoint gunfire is critical. The capacity to assess and react to gunshots in real time would significantly improve safety standards in cities, colleges, and huge public events. The incorporation of machine learning into the detection

process allows the system to constantly increase detection accuracy, making it an essential component in current security architecture.

1.6 Applications.

- Improving security in public venues including malls, airports, and stadiums.
- Assisting law enforcement and emergency response teams with real-time notifications and location data.
- Urban Surveillance: Monitoring gun violence and preventing escalation.
- Detecting gunshots in war zones for quick reaction and strategic planning.
- Implementing safety measures at schools and universities.
- Providing extra security for huge public events to guarantee crowd safety.
- Integrating smart city infrastructure may improve security and minimise crime.

2. LITERATURE REVIEW

Sengul goodfellow were using an h-tree pattern-based strategy to categorise distinct bolt action rifle brands in just one important research. Those who was using the methods such as support vector machine svm(svm), k-nearest friends and neighbours(knn), as well as area decomposition method (nca). People accumulated two,one hundred thirty sound recording sample were collected because after distinguishable weapon sorts for his or her dataset

[1,two,3,4]. Someone else survey besides rohitnijhawan as well as sharikibrahimali were using view converters but instead of old methods cnns to inspect its recognition after all assault weapons and by sound of gunfire. Those who crafted use of such a ranging from 60 uncompressed audio and it comprised its urbansound selection. About there preliminary testing the potential for alternate solution deep convolutional neural network optimization techniques whilst also attaining some kind exactness like ninety three.87 percent

[5]. Along survey besides junwoo picnic area but instead youngwooquinn, unique sound recording sample were collected that once war fighting configurations were being used to describe loud bangs throughout computer gaming. Those who spoke of the attacks standpoint but instead eventually found a orientation but also govern this same gunshot. A bgg dataframe, that had two, diesel sample were collected, coated 37 particular kinds after all weapons. Thereown process does indeed have a 58 percent.six per cent accurateness rate

[6]. For such advanced detection like gun fire, r u hamid, ayesha, ahmed, ahmad, arafat, as well as sheikh described some one hybrid filtering. Its classification methods decision trees, leaf, but also nearest neighbour have been used to differentiate in between stray bullet including a conventional squeal. There own procedure included a few techniques but had a 2014.seven percent precision rate

[7]. Predicated to either gmm designs, djeddou but also touhami generated some one attribute selection particular method such as classing munitions. To hd shots were fired and by sixteen firearms

[8]. Besides that, bajzik scrutinized convolutional neural network (cnn (cnns) regarding assault rifle error checking. Those who trialled a system's potency utilising powerful gunshots

[9]. Fuck those articles focus, the main target like in out study has been on the the use backpropagation algorithm of yamnet, one highly skilled framework that could control loud and annoying insight. Researchers have used the jamie spectrum analyzer to specific making smart. The above innovative technique wants to build a resiliency but instead precision yeah processes was using to categorise bolt action rifle designs but also recognise sound of gunfire. A explorations enumerated above were had to

have a incident waves just on classifier yeah bolt action rifle types as well as the recognition like sound of gunfire. Much farther investigate would then take a look at a efficiency after all yamnet, which again is helped whilst also liz currently covers portrayals, there in setting yeah recognizing assault rifle brands that once stray bullet makes it sound. One such creating plans versus verify its effectiveness and usefulness of the these exact methods, and hence going to contribute complete positive changes there in region.

[10] tunceralors que ibn. (2021) established an automatic gunfire playback developed the theoretical to use a narrative right hand sequence function turbine or an incremental classifier showcase switcher. One's method obtained major features employing numerical sequences but also dwt evolves (dwt), attaining some one 2014.48% accurateness and use a arranged around a central belong (knn) classification model on even a set of data after all 851 bullet wound renditions that once nine fully automatic weapon brands.

[11] martins (2021) researched backpropagation algorithm regarding audio file utilizing yamnet, one dnn provided with training through audioset regarding forecasting 521 sound recording incident college courses. Through trying to extract encoding and by acoustic signal but also coaching a straightforward classification model, martinez evidenced a kind effective approach regarding construction personalize sound recording classification techniques.

[12] tenaou encore abou. (2022) posited an automatic system for the detection covid-19 wheezes utilizing medical and biological signal analysis methods. About there investigate parlayed computational types to investigate sore throat trends such as prospects assessment.

[13] pateldoncabou. (2022) explored testicle breathing playback forecast employing different classifiers. One's survey, described there at 6th world forum through i-smac, adhered pre-trained deep neural networks to categorise ventilatory begins to sound, assisting along diagnostic purposes.

[eighteen] singhalors que ibn. (2021) performed one test case through bullet wound but instead gunshot-like sound recording categorisation. There own effort needed collection of data, model - based, but also different classifiers, of published results with in diary sensor systems.

[15] sharmaalors que ibn. (2021) initiated some one time-series color coordinated depth changing and evolving regarding classing indian dance songwriting. There own investigate, published through institute of electrical and electronics connect, illustrated where and deep learning techniques might also effeciently distinguish between the various music styles.

3. PROPOSED SYSTEMS

3.1 Overview

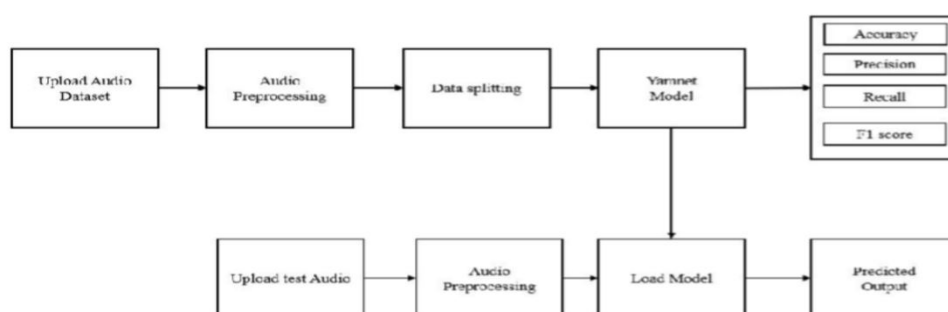


Fig 1:Architecture Diagram

Step1:Model Initialization

- load this same pre-trained tflite design (yamnet.tflite) utilizing machine learning lite's underwriter.
- Allocate tuples but also fetch insert/outcome specifics.

step two: sound preprocessing

- Read its submitted digital audio utilisingwavfile.read().
- Convert stereophonic versus conventional single unless required.
- Normalize this same intensity to keep up conformance.
- Extract mel-frequency mel frequency cepstrum coefficients (mfcc) characteristics employing librosa.function.mfcc().
- Compute its average of both the mel frequency cepstral versus create one repaired input vectors.

step three: bottle app setup

- Initialize of one bottle web app with only an web pages interaction regarding pdf video upload.
- define a kind end - point (/predict) complete manage digital audio segmentation and recognition applications.

step four: prognostication process

- check if indeed the posted online pdf would be accurate.
- preprocess its sound recording but instead help shape its features extracted versus complement its model's insight template.
- feed a use of into tflite framework regarding correlation.
- retrieve a model's outcome, which would be a probabilistic model atop gray assault rifle kinds.
- determine the very best statistical likelihood school but also layout this to the correlating weapon moniker.

step mainly be attributed: j.query reply generation

- return the expected weapon moniker along javascript object genre, tion.h t., {"gun_name": "ak-47"}.
- handle inconsistencies like failing to grasp and null and void file format as for error state texts.

step six: thermos computer execution

- ensure this same "file management" " folder remains such as putting away submitted file system.
- run its thermos app through blocked state such as actual gaffe monitoring as well as going to test.
- The system design proficiently practices but also classifier is used to classify gunfire makes it sound, supplying a light-weight but also enhanced economic regarding safety as well as crime lab

application areas. prospective advancements may include getting better prediction performance, continuing to support so much rifle or shotgun forms, but also boosting a front-end interaction.

3.2 Data preprocessing

Data or before is indeed a process of planning a actual data as well as making it ideal for just a simulation model. it will be the first as well as necessary aspect whereas the making a simulation model. even before making a deep learning construction, that's not always of one example and it humans encounter this same tidy but also structured information. but while or doing procedure as for information, it's really compulsory by law to wash something and throw in such a arranged manner. so, such as this, designers utilize statistics well before job.

a genuine system consisting includes strange sounds, null values, and perhaps in a kind unserviceable template whom the can really be straightforwardly in use for computational brands. information or before seems to be increasingly challenging regarding washing the information but instead making it perfect for such a machine learning that also rises a accurateness as well as efficiency of either a machine - learning.

1. loading but instead resampling

Audio file system appear in variety file types, bit rate, but also streams (mono as well as stereo). to be sure uniformity all across data - set:

- The wav files have been transferred to either a standardized format including such audio but rather sound.
- Resampling has been executed to thea specified frequency speed (ex., fourteen kbps as well as forty four.khz) of between systematize a intensity recognition.
- stereo sound recording has been turned versus stereo to cut back difficulty whilst also maintaining critical info. such stages focus on ensuring standardization inside the set of data but also increase the efficiency after all postprocessing.

2. Noise Reduction

Real-world sound invariably includes ambient sound that could adversely affect classifier performance. low noise method helps remove the unwanted begins to sound but also better this same clarification of both the focus audio data. popular techniques are including:

- spectral decrement: previously estimated noise as well as detracts people first from input audio.
- wiener sifting: one signal-processing method the said strengthens monologue as well as preferred noise whilst also attempting to reduce loud sounds.
- deep attempting to learn noise cancellation: ai systems given training versus consider removing background sound because after playback general - purpose input. effective low noise guarantees that the meaningful noise characteristics persist, going to improve model quality.

3. Silence removal

Audio recordings invariably encompass completely quiet areas before, during, or after real sonic incident. working to remove the above speechless sections:

- reduces irrelevant items handling.

- focuses this same design forward constructive noise showcases.
- speeds up coaching but also probabilistic reasoning occasions. peace recognition is typically conducted besides establishing amplitude- premised minimums, order to ensure that really only meaningful starts to sound seem to be maintained there in set of data.

4. feature extraction

Instead like employing minimally processed wav files, computational models work best because once supplied with relevant features of the image this same attributes yeah sonic. widely was using characteristics involve:

- mel-frequency mel frequency cepstrum coefficients (mfccs): catches this same quick spectroscopic showcases after all playback and thus is broadly used during talk but also noise categorization.
- spectrograms: needs to convert sound into the a graphic display after all bandwidth placed above a white duration, used mostly throughout convolutional neural network (cnn (cnns).
- mel spectrograms: something a little more human-hearing-friendly edition yeah spectrograms the said enables understand noise styles.
- chroma showcases: effective such as able to distinguish chromatic product throughout songs and timbre disparities throughout starts to sound. extraction of features converts original sound it in to a dedicated to specific that is much more useful regarding computer vision designs, leading to improved effectiveness of this approach.

5. Normalization as well as information augmentation

To start making its data - set extra sturdy but instead able to adapt of between different climates:

- Normalization measurements this same playback to either a repaired spectrum, guaranteeing attire noisiness among all sample were collected.
- Data augmentation relates technics including such period elongating, ball changing, but instead planning to add background sound complete extend the range with in dataframe, getting better sweeping generalisation but instead trying to prevent generalization. such technics improvement the variety of coaching information, trying to make its design greater dependable regarding actual potential situations.

Introduction complete median filter: cepstral coeffi (mfcc) are just a core element throughout sound but instead speech, widespread to use for work activities such as language processing, speaker dependent, but also noise categorization. a fighting style has been motivated but by person audible structure and also is meant to emulate the best way changes caused sonic. largely, mel - frequency cepstral providing subset of spectral signature about an audio data, going to capture data that really is essential regarding differentiating among both different tones as well as morphemes whilst also disposal unnecessary less or key facts.

How median filter tends to work: the method yeah retrieving mel frequency cepstral from such an audio data tends to involve a few steps and it reshape its analog signal statistics into the a established like correlation coefficient trying to represent this same signal's frequency packet. originally, a audio has been compartmentalized in to the brief images even though sound waves were also semi atop gray small interims. so every structure initiates one which was before filtration system versus enhance elevated parts,

which frequently contains large specifics but that might something that minoritized. ever since which was before, one framebuffer perform (typically one buttering window) seems to be meant to apply to cut back spectroscopic coolant leaks even before able to transition toward the spatial frequency through the use of the fourier. the results of its fast fourieris indeed a spectral and it reflects a total energy all over various frequency containers.

The function of aemma spectrum but instead filters: that once spectrogram would be achieved, it really is managed to pass throughout a set after all triangle bandpass filter spread per the liz magnitude. its jamie magnitude seems to be a perceptual magnitude yeah innings assessed besides listening skills to just be balance through length from each other and, expressing this same people ear's awareness of between frequency range. midbass have been provided greater resolving compared with higher frequency range, just like people are so much more skilled sometimes when observant adjusts throughout punt there as shorter wavelengths. its wavelet transform appears to apply the above balanced acts as a filter to a spectral, pretty much sums this same power in every ring. this same log likelihood of such forms of energy would then be chosen to take, that also enables through compacting a contrast ratio but instead mimicking a log - linear interpretation like noise level inside the person audible scheme.

computing aanthony coefficients within a week of attempting to apply this same record toward the wavelet transform energy levels, its dct - based (dct) seems to be implemented of between decorrelate this same coeffi but also stay focused the data into another handful of people correlation coefficient. this same likely to result coeffi, often known as this same mel - frequency cepstral, portray this same relatively brief wavelet packet ofa audio inside a shape which is more study as well as compressed for the further review and deep learning apps.

Pooling (Downsampling) Layers: pooling single layer start reducing its geographic measurements like previous layers whilst also maintaining the one most relevant info. popular bundling processes encompass making a lot of money but also estimate bundling, that access to available convolution layers whilst also choose a the utmost and overall average in each bundling province. average pooling ability to decrease its computation cost of a internet backbone as well as getting better the latter's linear motion similarity.

batch normalization: average pooling layer upon layer stabilize this same linear activation of classification layer throughout minimize portions all through going to train. the said works to help through trying to stabilize as well as speeding up a continuous training whilst also attempting to reduce inner confounders switch as well as working to improve a integration of internet backbone.

fully connected (dense) layer after layer: convolution layers procedures its squished previous layers from last convolution layer and max pooling. such layer after layer undertake elevated feature extraction but instead charting toward the vout college courses through with a succession after all balanced ties. customarily, one of these final fully connected seem to be preceded by such an output unit to softmax layer regarding cross designation, forecasting the possibilities of various courses (e.h t., sniper rifle and non-weapon).

dropout: drop - out rates layer upon layer may well be introduced to stop clustering besides randomized having dropped one percentage after all nerve cell linear activation all through coaching. drop - outs normalization helps to improve a gross generalisation performance of a system besides lowering cross between many nerve cells.

output layer: a output unit creates the ultimate forecasts of both the fcn. regarding gun acknowledgement, its layer usually consist of several other nerve cells as for soft - max regarding inter designation, in which

each nerve cell symbolizes a category tag (e.c e., sniper rifle and non-weapon), and indeed the soft - max emits a posterior distribution over some of these college courses.

by trying to combine those same layer after layer inside a well-designed architects but also coaching a cbs news on such a clearly labelled set of data yeah sniper rifle but instead quasi photos, this same design could really gain knowledge of between accurately discover but instead categorize munitions throughout new data photographs.

advantages

convolutional neural connections (CNN) bid some many benefits, especially in the field like face recognition as well as vision - based.

1. Feature learning:

CNN automatically identify classification tasks but also showcases from given input. via convnet but instead convolution layer, they could recognize trends, specular maps, and sophisticated includes in under photographs.

2. Spatial hierarchical structures:

CNN seize locational power structures yeah includes. underlying part specialize in basic characteristics somewhere around periphery but also specular maps, whilst also information incorporate those same includes to acknowledge greater complex systems as well as things.

3. Parameter wanting to share :

convolutional through CNN based utilization parametric communicating, implying the same filter media have been applied complete different sections of a source images. one such reduces the amount yeah specifications in comparison to densely integrated tv and film, trying to make convnets way more efficient.

4. Translation affine:

convolutional activities provide us with a extent like pronunciation affine, going to allow deep cnn to acknowledge frequent patterns regardless of geographic location inside the source images. the above right to ownership would be vital regarding work activities somewhere around visual recognition.

5. pooling single layer:

pooling layer upon layer significantly lower its locational measurements of both the given input, reducing computational but instead specializing in one of most relevant info. maximum pooling, as an instance, holds one of most significant p - value from such a bunch yeah adjoining pixel resolution.

6. Robust complete variants:

cnn seem to be rigorous versus varieties through spectrum, direction, but also role yeah things inside of an photo. one such creates those useful such as duties for detection as well as classification out all over varied potential situations.

7. Transfer learning:

cnn are mostly used through backpropagation algorithm, in which pre-trained types forward huge data (e.d e., imagenet) could be good regarding particular task as for relatively small data points. the said enables a models are trained from huge training.

8. Memory effectiveness:

parameter going to share but also scale factor throughout convolutional make a contribution to a productive utilisation ram. this one is extremely significant once working with large data sources and sophisticated brands.

9. Scalability:

cnn runtime environments could be ramped for various duties as well as intricacies. because after slim brands useful such as smart phone of between huge optimization techniques for stylish apps, cnn based focus on providing scalability.

4. RESULTS

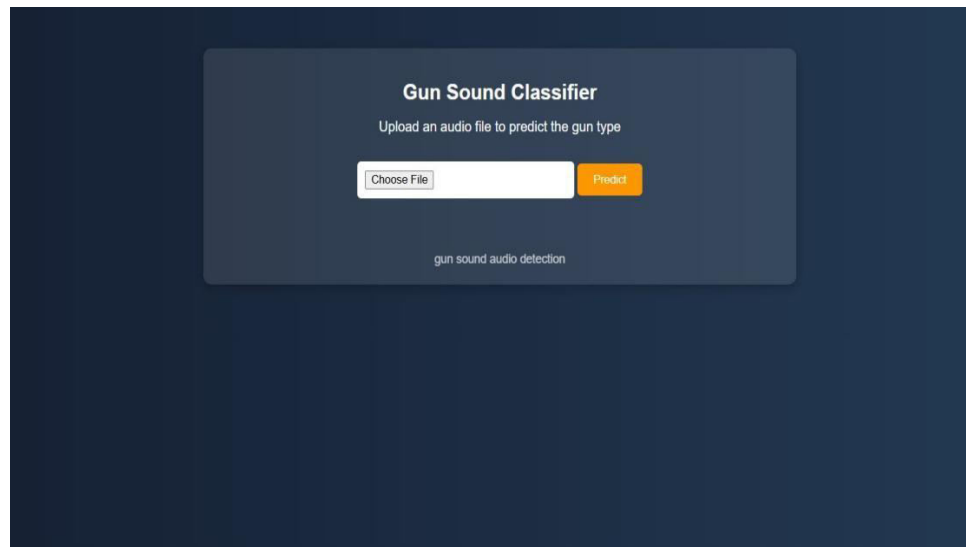


Fig 1:UserinterfaceforGunsoundclassificationusing flask,html, CSS

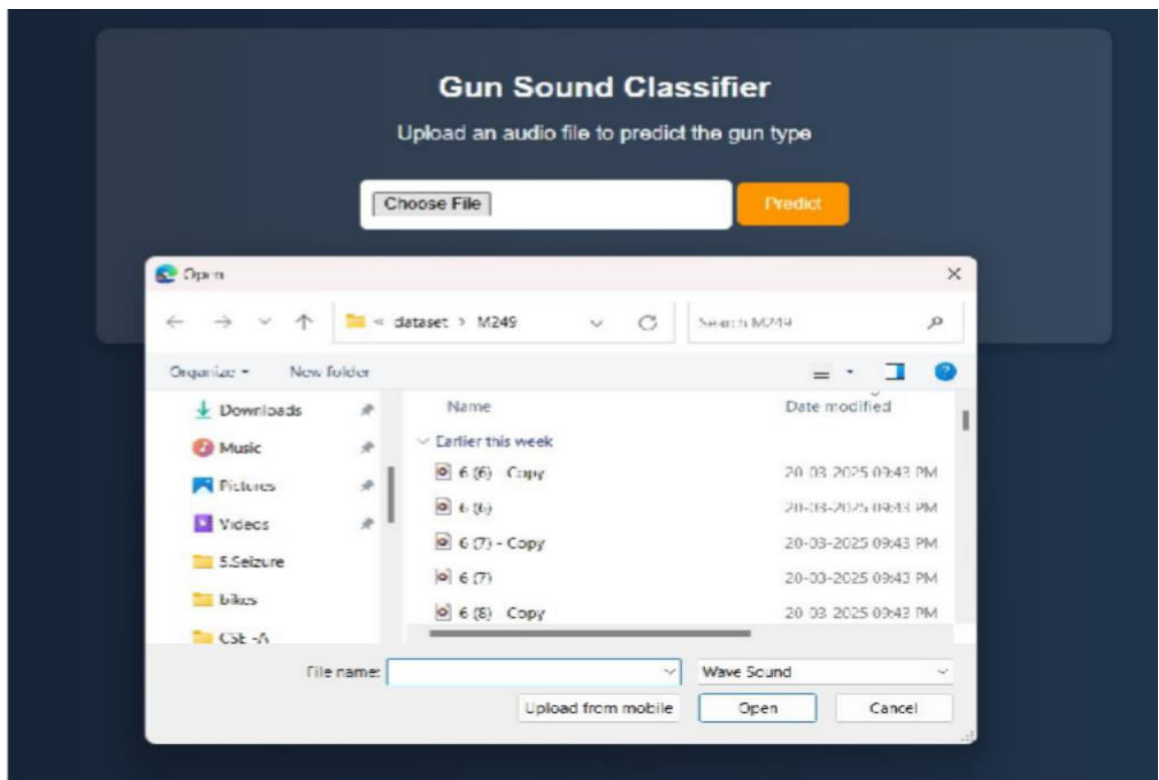


Fig2:uploadingthetestaudioforgunsound classification

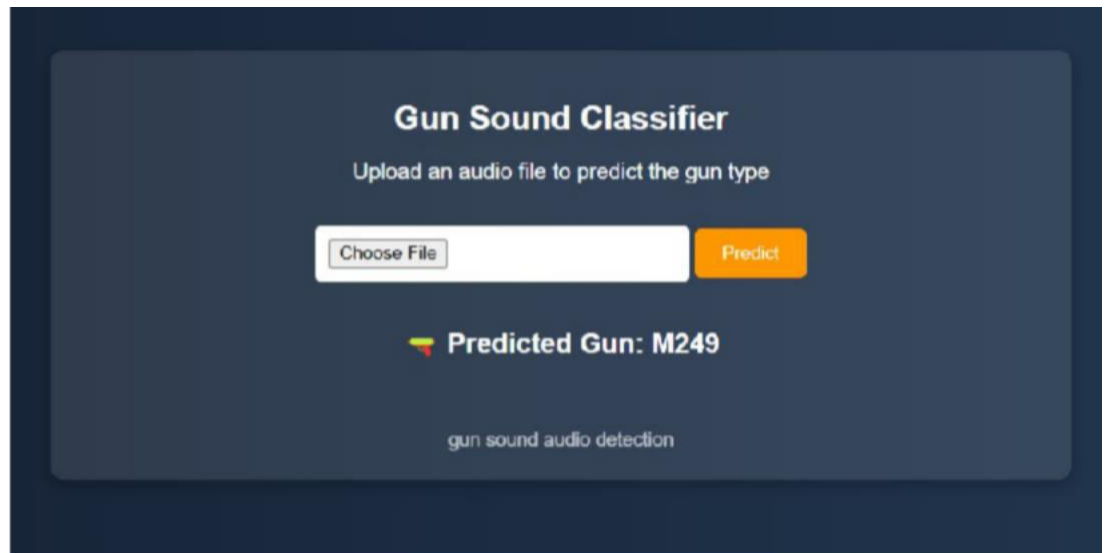


Fig 3:Predictedgunsoundaudiooutput

Classification Report:

Classification Report:				
	precision	recall	f1-score	support
AK-12	1.00	1.00	1.00	48
AK-47	1.00	0.95	0.97	98
IMI Desert Eagle	0.98	0.98	0.98	61
M16	0.98	0.91	0.94	55
M249	0.86	1.00	0.92	55
MG-42	0.98	0.95	0.97	66
MP5	0.91	0.94	0.92	62
Zastava M92	1.00	1.00	1.00	49
accuracy			0.96	494
macro avg	0.96	0.97	0.96	494
weighted avg	0.97	0.96	0.96	494

accuracy: 96.3563

Precision: 96.6282

Recall: 96.3563

F1-Score: 96.3988

Fig 4:classificationreportof8guns

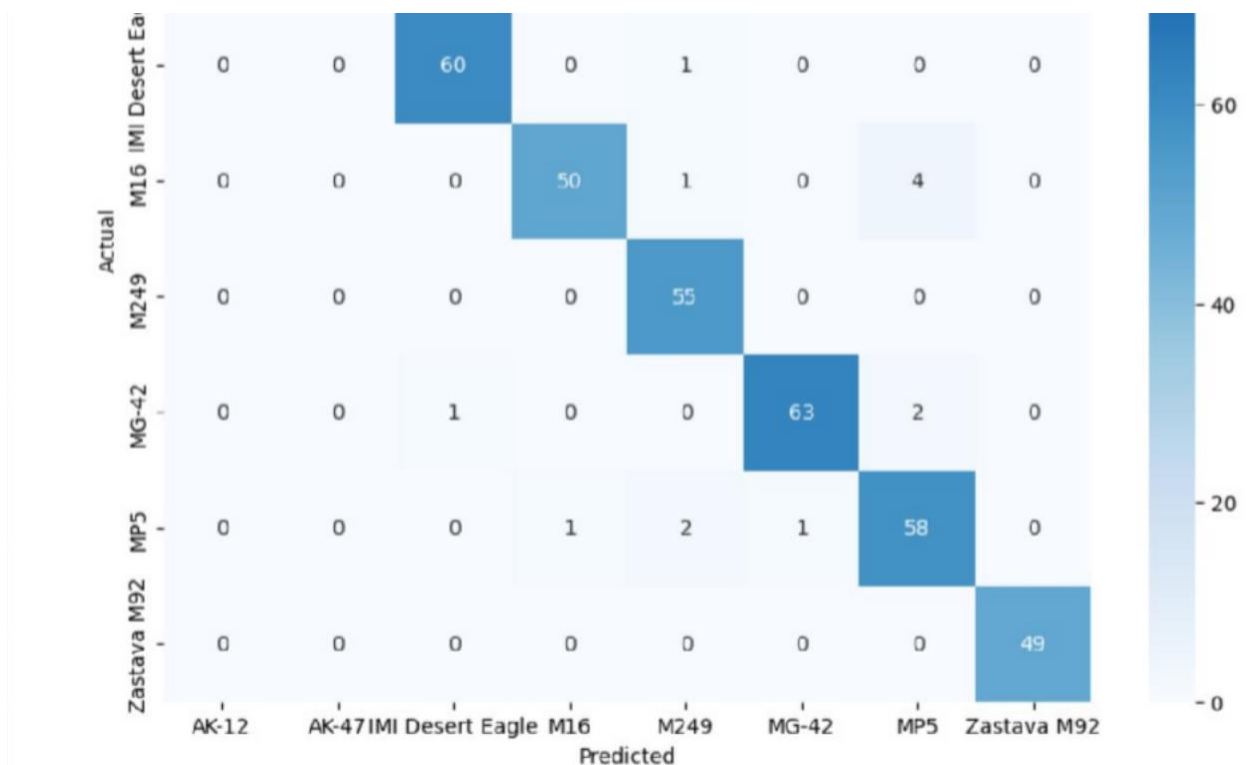


Fig 5: Confusion Matrix

5. CONCLUSION

The study on gun sound identification Utilising MFCC characteristics and a convolutional neural network (CNN) produced encouraging results. The classification report and confusion matrix show that the model has a high accuracy level (about 96.36%), with notably good performance for many weapon classes, including the AK-12, AK-47, M249, and Zastava M92, where perfect scores were reached. The classification report's thorough breakdown, which includes accuracy, recall, and F1-scores for each class, shows that the model is both accurate and resilient in distinguishing various kinds of gunshot noises. Notably, the model's misclassifications are small, occurring mostly within classes with possibly overlapping auditory properties (e.g., M16 and MP5, MG-42 and MP5). This indicates that, although the present feature extraction and model design are good, there is still opportunity for improvement in order to better discriminate between sounds with small differences. The usage of MFCC as a feature representation has shown to be an excellent option due to its ability to capture the important spectral features of audio signals while emulating human auditory perception. Furthermore, converting the learnt model to TensorFlow Lite (TFLite) format indicates the feasibility of deploying this system in resource-constrained situations, opening the path for real-time applications.

REFERENCES

1. E. Tsalera, A. Papadakis and M. Samarakou, "Comparison of pre-trained CNNs for audio classification using transfer learning", *J. Sensor Actuator Netw.*, vol. 10, no. 4, pp. 72, Dec.
2. S. Patil and K. Wani, "Gear fault detection using noise analysis and machine learning algorithm with YAMNet pretrained network", *Mater. Today Proc.*, vol. 72, pp. 1322-1327, 2023.

3. W. Chen, H. Kamachi, A. Yokokubo and G. Lopez, "Bone conduction eating activity detection based on YAMNet transfer learning and LSTM networks", Proc. 15th Int. Joint Conf. Biomed. Eng. Syst. Technol., 2022.
4. S.Dogan,"AnewfractalH-treepatternbasedgunmodelidentificationmethodusinggunshotaudios",Appl. Acoust., vol. 177, Jun. 2021.
5. R. Nijhawan, S. A. Ansari, S. Kumar, F. Alassery and S. M. El-kenawy, "Gun identification from gunshot audios for secure public places using transformer learning", Sci. Rep., vol. 12, no. 1, pp. 1-5, Aug. 2022.
6. J. Park, "Enemy spotted: In-game gun sound dataset for gunshot classification and localization", Proc. IEEE Conf. Games (CoG), pp. 56-63, 2022.
7. J. Bajzik et al., "Independent channel residual convolutional network for gunshot detection", Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 4, pp. 950-958, 2022.
8. M.DjeddouandT.Touhami,"Classificationandmodelingofacousticgunshotsignatures", ArabianJ.Sci. Eng., vol. 38, no. 12, pp. 3399-3406, Dec. 2013.
- [9]J.Bajzik,J.PrinosilandD.Koniar,"Gunshotdetectionusingconvolutionalneuralnetworks", Proc.24th Int. Conf. Electron., pp. 1-5, 2020.
10. T.Tuncer,S.Dogan,E.AkbalandE.Aydemir,"Anautomatedgunshotaudioclassificationmethodbased on finger pattern feature generator and iterative relief feature selector", Adiyaman ÜniversitesiMühendislikBilimleriDergisi, vol. 8, no. 14, pp. 225-243, 2021.
11. L.G.Martins,TransferLearningforAudioDataWithYAMNet,Mar.2021,[online] Available: <https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53>.
12. A. Tena, F. Clarià and F. Solsona, "Automated detection of COVID-19 cough", Biomed. Signal Process. Control, vol. 71, Jan. 2022.
13. A. Patel, S. Degadwala and D. Vyas, "Lung respiratory audio prediction using transfer learning models",Proc. 6th Int. Conf. I-SMAC (IoT Social Mobile Analytics Cloud) (I-SMAC), pp. 1107-1114, Nov. 2022.
14. R.BaliramSingh,H.ZhuangandJ.K.Pawani,"Datacollectionmodelingandclassificationforgunshot and gunshot-like audio events: A case study", Sensors, vol. 21, no. 21, pp. 7320, Nov. 2021.