Analysis of Voice Changes in Anti Forensic Activities Case Study: Voice Changer with Telephone Effect

Abiyan Bagus Baskoro[1], Niken Dwi Cahyani[1*], Aji Gautama Putrada [1]

[1]School of Computing, Telkom University
Bandung, Indonesia

* nikencahyani@telkomuniversity.ac.id,

Abstract

Voice recordings can be changed in various ways, either intentionally or unintentionally, one of which is by using a voice changer. Reference voice recordings and suspect voice recordings will be more difficult to analyze if suspect voice recordings are changed using a voice changer application under certain effects such as telephone effect. Voice Changer can be one form of activity that can be carried out by anti-forensics, making it difficult for investigators to investigate if the voice recording is changed with telephone effect. This study has two types of recordings, namely the reference voice recording (unknown sample) and suspect voice recording (known sample) that has been changed using a voice changer application with telephone effect. Investigations were carried out based on data results extraction and analysis using pitch, formant, and spectrogram using the Analysis of variance (ANOVA) method and the likelihood ratio method. The results of this study indicate that the application of voice changer can be one form of activity that can be carried out by anti-forensics so that it can be difficult for investigators to conduct investigations on sound recording evidence. This research may help forensic communities, especially investigators to conduct investigations on sound recording.

.

Keywords: voice changer, analysis of variance (ANOVA), likelihood ratio

I.   INTRODUCTION

The term anti-forensics becomes the basic term for digital researchers. Although this term is not something new, it does not have a clear set of definitions [1]. Forensics is a specific scientific analysis of anti-forensic behavior as evidence presented in court. Since 2008, sound recordings can be accepted as legal evidence in Indonesian courts [2]. With the existence of multimedia technology to produce sound recordings, evidence is often found in the form of a sound recording tool at the scene of a case where there is a recorded voice of someone's conversation. Different voice recorders found in crime scenes can be compared to determine whether the recording device is from the same person or not [3] .This digital audio recording is a piece of evidence that needs to be verified for authenticity, considering that its use in court proceedings is quite high and continues to increase [4].   Forensic Voice Comparison (FVC) is usually concerned with the comparison of offender recordings with suspect recordings, with the aim of assisting investigative authorities or the court in identifying the identity of a speaker [5]. Records can be presented in criminal and civil cases as evidence. Voice recordings

made of evidence (unknown samples) can be changed intentionally or unintentionally by some parties. The sound recording with changes in the structure of audio or file metadata file could be called as container-based, while addition, hiding or cutting voices can be called as content-based tampering[6] .Voice changes can be done using several methods; a voice changer application is one of them. In a voice changer application, several methods can be used to change evidence (unknown samples) or references (known samples) by changing the pitch, reverb, tone, or changing the sound into several options; for example, robotic voices, human voices, or radio sounds making it difficult for investigators to investigate voice recordings. The words of each recording comparison are considered identical if they meet each of the predetermined parameter requirements. Otherwise, the words from each recording comparison are considered not identical if they do not meet each of the predetermined parameter requirements.

A previous study have been conducted in the testing of voice recordings using voice changers by changing the pitch of sound recordings to high pitch and low pitch [7]. The results of the study concluded that sound recordings that have been changed to low pitch have a higher level of identification when compared to high pitch, by utilizing pitch, formant and spectrogram analysis. To the extent of this study, our current research aims to prove that whether the application of sound modifiers can be a form of activity that can be carried out by anti-forensics that can make it difficult for investigators to conduct an investigation. One of which is when the voice recording is changed with telephone effect. Our analysis is based on comparison of data integrity in voice recordings and analyzing the voice recording using pitch, formant and spectrogram analysis with the ANOVA method and the likelihood ratio. Result of this research can be used as a basis for the investigator, to determine whether the voice recordings are changed using voice changer can still be identified or not by the forensic community to investigate the voice recording

II. LITERATURE REVIEW

A. Audio Forensics

Audio forensics is the science of processing digital evidence consisting of sound recordings of an act of crime. The sound recording must be analyzed to determine and verify the identity of the speaker in the recording [8]. Recognition of the identity of a speaker is known as voice recognition. The technique used is by comparing the sound of digital evidence (unknown samples) with recognized sounds (known samples). If there is an identical match between the unknown sample and the known sample, then it can be concluded that the evidence is really the sound of the known sample. The Standard Operation Procedure (SOP) on Audio Forensic Analysis from the Digital Forensic Analyst Team (DFAT) Digital Forensic Laboratory Centre at Bareskrim POLRI, states that in doing a sound recording analysis, the steps are acquisition, audio enhancement, noise filtering, decoding and voice recognition.

The acquisition process is the process of acquiring digital evidence, by recording all information related to digital evidence from storage media to produce an image file of the evidence. The process of obtaining an image of evidence must be verified by comparing the original file hash with the image file hash. In audio forensics, investigators need to get reference voices (known samples) from suspects. The suspect voice recording process must be with a video that shows the subject is talking. The voice source must be clear and attached with an official agreement and a memorandum of agreement signed by the subject whose voice will be analyzed.

The enhancement and noise filter process is a stage of improving the quality of sound recordings. This stage needs to be done if the sound recording to be analyzed has a fairly high noise level [7]. However, if the sound recording quality is good, then this process does not need to be done again.

Decoding is the process of making a voice recording transcript and is carried out by at least two examiners. Record transcripts must include the subject label (e.g. subject 1, subject 2, and so on) and the time (in hours: minutes: seconds) that corresponds to the recording. If the voice of the conversation on the recording sounds unclear, it is written "unclear" so that the results of the transcript only contain clear and understandable speech pronunciation of words.

Meanwhile, voice recognition process is carried out to ensure whether the sound on the recorded evidence is identical to the suspect voice or not. In this process, analysis of pitch, formant, and spectrogram is carried out. A minimum of 20 exact words between the evidence recording and the suspect recording is required.

B. Anti-forensics

Anti-forensics is a method used to challenge forensics investigators in investigations. The purpose of this method is to hide, modify, or destroy digital evidence so that investigators have difficulty investigating digital evidence and result in the digital evidence being unacceptable in legal proceedings [9]. In the book Anti Forensics by Sulianta, it is stated that anti-forensic objectives generally include two things [10]. First, how to make data that cannot be found or opened, for example by hidden, encryption, and steganography techniques; Second, is how to ensure that if found, the data will still not be in accordance with legal standards. Anti-forensics is a method or attempt to hide traces of tampering so that it can escape the forensic method.[11].

With the existence of multimedia technology to produce sound recordings, the evidence is often found in the form of a sound recording tool at the scene of a case where there is a recorded voice of someone's conversation. Evidence of the sound recording tool can contain a someone's conversation about the crime he has committed. Data can be manipulated by creating a "faked" version with the purpose of appearing to be something else [12]. If the investigator compares the voice on the evidence with the voice on the suspect, the suspect will not be found guilty because the voice pronounced by the suspect is different from the voice on the voice recorded evidence.

There are several methods to change the sound recording of evidence, voice changer application is one of them. The voice changer application can change the sounds of sound recording evidence into several options by changing the pitch, reverb, tone, or changing the voice recording into several options; for example, robotic voices, human voices, radio, or telephone. In order to make the investigator unsuspecting about the changes in the voice recordings, the voice recordings can be changed to telephone conditions, because the original sounds on the evidence and the voice which have been changed into telephone conditions can still be said to be the same or not much different with the original

C. Statistical Analysis of Pitch

Pitch is a basic frequency of vibration results from the vocal cords that are used for speech signal processing [13]. Each person has a different and unique pitch, where it is influenced by the physiological aspects of each person's larynx. Under normal speech conditions, men have pitch levels ranging from 50 to 250 Hz and women have pitch levels ranging from 120 to 500 Hz [14]. As pitch is unique for each person, pitch analysis can be used as voice recognition through statistical analysis of the minimum pitch, maximum pitch and mean pitch values. Pitch analysis is done based on the calculation of the difference in pitch value of each voice changer and reference recording. The pitches of each recording are compared based on the minimum, maximum, and mean value of each pitch.

D. Formant Statistical Analysis

At the stage of formant statistical analysis, it is necessary to do 2 types of analysis, namely:

1) Analysis of Variance (ANOVA)

Formant is the resonant frequencies of a filter. The vocal tract (articular) will continue to filter the periodic sound from the vibrations of the vocal cords into the output sound in the form of words that have meaning. In general, Formant frequencies are unlimited, generally drawn from Formant 1 (Frm1) to Formant 5 (Frm5), but to identify someone's voice, at least 3 (three) Formants are analyzed, namely Formant 1 (Frm1), Formant 2 (Frm2), Formant 3 (Frm3) [15]. Given recent technological advances in telephony (e.g. WeChat or WhatsApp) higher formants (Frm4-Frm5) are becoming increasingly part of evidence material [16]

2) Likelihood Ratio Analysis

A more detailed review of the formant analysis and bandwidth statistical analysis is by using the Likelihood Ratio (LR) formula as in (1).

$$LR = \frac{p\,(\,E\,|\,Hp\,)}{p\,(\,E\,|\,Hd\,)} \tag{1}$$

In this formula, we can assume that p(E|Hp) is a prosecution hypothesis, it means Known and Unknown Samples are from the same person. p(E|Hd) is a defense hypothesis, it means Known and Unknown Samples come from different people. p(E|Hp) comes from the ANOVA p-value, and p(E|Hd) = 1 - p(E|Hp).

If LR > 1, then this supports p(E|Hp), conversely if LR < 1, then p(E|Hd) is supported. For this reason, p(E|Hp)> 0.5 must be satisfied to be able to conclude that the sound of evidence (Unknown) and the sound of suspect (Known) are from the same person (Identical). The amount of the LR is followed by a verbal statement to explain the value of the LR, as in Table I and Table II [17].

Based on these two tables, to get support for the prosecution hypothesis (Unknown and Known voices are from the same person), the Likelihood Ratio needs to be more than 1. The greater the Likelihood Ratio value, the better and stronger the Verbal Statement.

E. Spectrogram

The spectrogram is a basic visual representation of the color of the sound (spectral) which varies over time[18]. The variation shows the level of energy intensity of the spectral. In other words, a spectrogram is a form of visualization of each Formant value which is equipped with an energy level that varies at different time. A spectrogram contains things that are detailed and sometimes it is also known as voice fingerprints, as, in the pronunciation of words and special pattern, the spectrogram forms a general pattern that is distinctive for each Formant value. Therefore, with the typical general pattern, the spectrogram can be analyzed to identify someone's voice [19].

TABLE I
VERBAL STATEMENT FOR EACH LR VALUE THAT SUPPORTS THE PROSECUTION HYPOTHESIS P(E|HP) (Adapted from Al-Azhar, 2010, p.224)

LR	Verbal Statement	Information
>10.000	Highly Strong Evidence to Support	
1.000 – 10.000	Strong Evidence to Support	Supports the
100 – 1.000	Moderately Strong Evidence to Support	prosecution hypothesis
10 – 100	Slightly Evidence to Support	$p(E|H_P)$
1-10	Limited Evidence to Support	

TABLE II
VERBAL STATEMENT FOR EACH LR VALUE THAT SUPPORTS THE DEFENSE HYPOTHESIS P (E | HD) (Adapted from Al-Azhar, 2010, p.224)

LR	Verbal Statement	Information
1 – 0,1	Highly Strong Evidence to Support	
0,1 – 0,01	Strong Evidence to Support	Supports the defense
0,01 – 0,001	Moderately Strong Evidence to Support	hypothesis
0,001 – 0,0001	Slightly Evidence to Support	$p (E | H_d)$
<0,0001	Limited Evidence to Support	

III. RESEARCH METHOD

A. Workflow

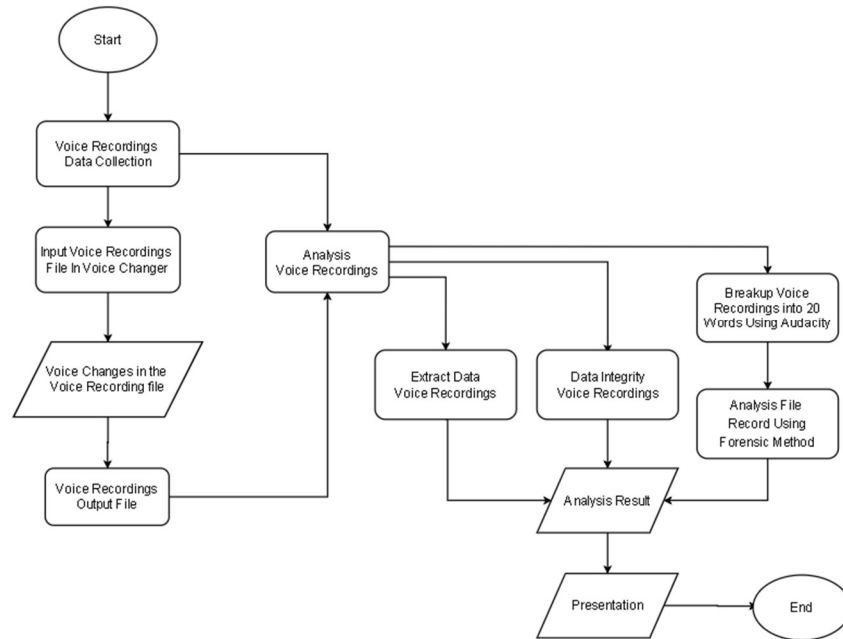The workflow of sound recording analysis is shown in Fig. 1.

Fig. 1 Sound Recording Analysis Flow

There are two phases in testing. The first phase is to prove that the standard operation procedure (SOP) on Audio Forensic Analysis from the Digital Forensic Analyst Team (DFAT) Digital Forensic Laboratory Centre at Bareskrim POLRI can be used as an Audio Forensic procedure. The second phase is to prove that a voice changer with telephone effect can be used as a tool for Anti-Forensics.

For the first phase, two records from every 5 people of each gender, male and female, will be used for proof. The first records use Samsung Galaxy S8 and will be used as the recognized voices (known samples) or reference. The second records use Samsung Galaxy A7 and will be used as evidence voices (unknown samples) or suspect voices. Both sound recordings are then processed through the audio-forensic procedure that is being tested. If the number of identical voices is bigger than the number of non-identical voices, the Audio Forensic method will be considered as proven.

For the second phase, two records from every 5 people of each gender, male and female, will be used for proof. The first records use Samsung Galaxy S8 and then will go through the process of voice changer with telephone effect. Then the output of the process will be used as the recognized voices (known samples) or reference. The second recording will be used as evidence voices (unknown samples) or suspect voices. Both sound recordings are then processed through the audio-forensic procedure that has been proven. If the number of non-identical voices is bigger than the number of identical voices, then the Anti-Forensic method will be considered as proven.

B. Hardware Specification

The specification of hardware used in this study is presented in Table III.

TABLE III
TEST ENVIRONMENT SPECIFICATION

Voice Changer Application		
No.	**Specification**	**Detail**
1	Name	Voice Changer – Audio Effects
2	Version Name	1.5.7
3	Developer/Publisher	Super Effect Studio
4	Sdk Version	21
Recognized Voice Recording Mobile Hardware		
No.	**Specification**	**Detail**
1	Name	Samsung Galaxy S8
2	Model	SM-G950FD
3	OS Version	Android 9.0 "Pie"
4	Chipset	Exynos 8895
5	Baseband Version	G950FXXS8DTC1
6	Internal	64GB
7	RAM	4GB
Evidence Voice Recording Mobile Hardware		
No.	**Specification**	**Detail**
1	Name	Samsung Galaxy A7
2	Model	SM-A720F/DS
3	OS Version	Android 8.0 "Oreo"
4	Chipset	Exynos 7880
5	Baseband Version	A720FXXU9CSJ3
6	Internal	32GB
7	RAM	3GB
Personal Computer Hardware		
No.	**Hardware**	**Detail**
1	Type	ROG GL503GE
2	CPU	Intel® Core™ i7-8750H Processor 2.2 GHz
3	IGPU	Intel UHD Graphics 630
4	RAM	16 GB DDR4, 2666 Hz
5	Graphics	NVIDIA GeForce GTX1050Ti 4GB
6	Operation System	Windows 10 Home

Samsung Galaxy S8 is used as the recognized voice recording device and the specification can be seen in Table III. Additionally, Samsung Galaxy S8 is used as the voice changer application. Samsung A7 is used as the evidence voice recording and the specification can be seen in Table III. The data of all original sound recordings will be analyzed using a laptop with specifications that can be seen in the Table III.

C. Software Specification

The specification of software and dataset used in this study is presented in Table IV. As seen in Table IV, Audacity is used to split the voice recording into 20 words. Android Debug Bridge (ADB) is used to extract voice recording from mobile hardware. USB Blocker is used to maintain the integrity of the extracted data. Praat and Gnumeric spreadsheet are used to analyze each word from voice recording. The dataset used in this study can also be seen in the table IV. The dataset obtained from mobile hardware (recognized voice and evidence voice) with five males and five females voice recordings from the same person each. Each sound recording file has mp3 file type.

TABLE IV
TEST ENVIRONMENT SPECIFICATION

Software

No.	Software	Detail
1	Tools	- Audacity
		- Praat
		- Android Debug Bridge (ADB)
		- Gnumeric Spreadsheet
		- USB Blocker

Dataset

No.	Type	Detail
1	Audio	20 files voice record
		- 5 male voices (.mp3) from Samsung S8
		- 5 female voices (.mp3) from Samsung S8
		- 5 male voices (.mp3) from Samsung A7
		- 5 female voices (.mp3) from Samsung A7

IV. RESULTS AND DISCUSSION

A. Data Extraction

The file directory of the evidence voice recordings on the Samsung Galaxy S8 is stored in the /sdcard/Recorders directory. Meanwhile the file that has been changed using the Voice Changer application is stored in the /sdcard/VoiceChangerStudio directory. The file directory of the evidence voice recordings on the Samsung Galaxy A7 is stored in the /storage/VoiceRecorder. After extracting the data, the recorded file is stored in the D:\ADB Extract directory.
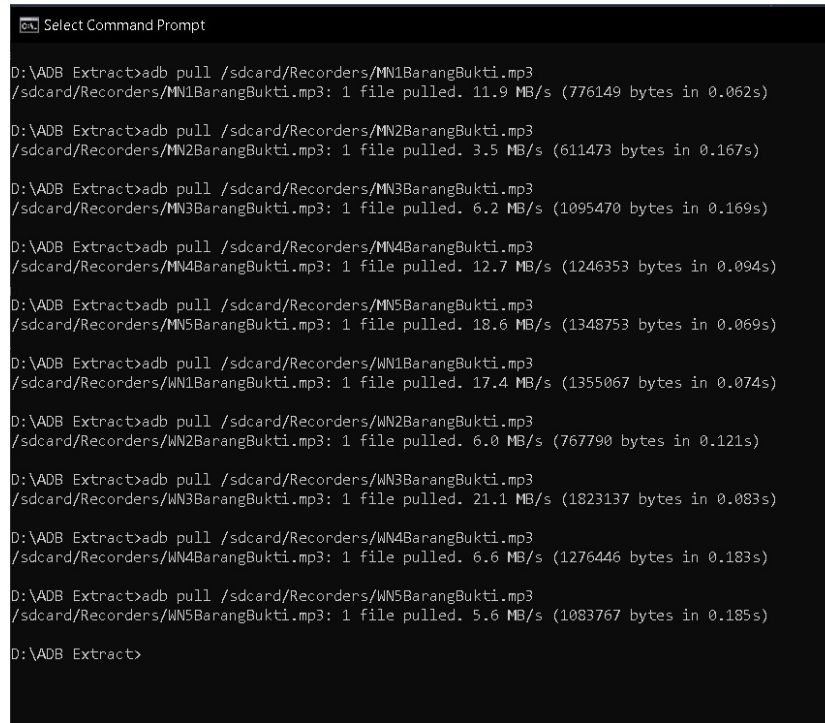
The recognized voice consists of five male voices and five female voices. The five male voice recordings are given the code MN 1, MN 2, MN 3, MN 4, MN 5. The five female voice recordings are given the code WN 1, WN 2, WN 3, WN 4, and WN 5. All recordings are in mp3 file format. Process of recognized voice recording extraction can be seen in Fig. 2.

Fig. 2. Recognizing voice recording extraction

The evidence voice consists of five male voices and five female voices. The five male voice recordings are given the code MN 1', MN 2', MN 3', MN 4', MN 5'. Notice the quotation marks to distinguish these records from the recognized records. The five female voice recordings are given the code WN 1', WN 2', WN 3', WN 4', and WN 5'. The female codes also have quotation marks to distinguish them from the recognized records. All recordings are in mp3 file format. Process of evidence voice recording extraction can be seen in Fig. 3.


```
ox. Select Command Prompt

D:\ADB Extract>adb pull /sdcard/Recorders/MN1BarangBukti.mp3
/sdcard/Recorders/MN1BarangBukti.mp3: 1 file pulled. 11.9 MB/s (776149 bytes in 0.062s)

D:\ADB Extract>adb pull /sdcard/Recorders/MN2BarangBukti.mp3
/sdcard/Recorders/MN2BarangBukti.mp3: 1 file pulled. 3.5 MB/s (611473 bytes in 0.167s)

D:\ADB Extract>adb pull /sdcard/Recorders/MN3BarangBukti.mp3
/sdcard/Recorders/MN3BarangBukti.mp3: 1 file pulled. 6.2 MB/s (1095470 bytes in 0.169s)

D:\ADB Extract>adb pull /sdcard/Recorders/MN4BarangBukti.mp3
/sdcard/Recorders/MN4BarangBukti.mp3: 1 file pulled. 12.7 MB/s (1246353 bytes in 0.094s)

D:\ADB Extract>adb pull /sdcard/Recorders/MN5BarangBukti.mp3
/sdcard/Recorders/MN5BarangBukti.mp3: 1 file pulled. 18.6 MB/s (1348753 bytes in 0.069s)

D:\ADB Extract>adb pull /sdcard/Recorders/WN1BarangBukti.mp3
/sdcard/Recorders/WN1BarangBukti.mp3: 1 file pulled. 17.4 MB/s (1355067 bytes in 0.074s)

D:\ADB Extract>adb pull /sdcard/Recorders/WN2BarangBukti.mp3
/sdcard/Recorders/WN2BarangBukti.mp3: 1 file pulled. 6.0 MB/s (767790 bytes in 0.121s)

D:\ADB Extract>adb pull /sdcard/Recorders/WN3BarangBukti.mp3
/sdcard/Recorders/WN3BarangBukti.mp3: 1 file pulled. 21.1 MB/s (1823137 bytes in 0.083s)

D:\ADB Extract>adb pull /sdcard/Recorders/WN4BarangBukti.mp3
/sdcard/Recorders/WN4BarangBukti.mp3: 1 file pulled. 6.6 MB/s (1276446 bytes in 0.183s)

D:\ADB Extract>adb pull /sdcard/Recorders/WN5BarangBukti.mp3
/sdcard/Recorders/WN5BarangBukti.mp3: 1 file pulled. 5.6 MB/s (1083767 bytes in 0.185s)

D:\ADB Extract>
```

Fig. 3. Evidence of voice recording extraction

The voice changed recognized voice consists of five male voices and five female voices. The five male voice recordings are given the code MN 1, MN 2, MN 3, MN 4, MN 5. The five female voice recordings are given the code WN 1, WN 2, WN 3, WN 4, and WN 5. All recordings are in mp3 file format. Process of voice changed recognized voice recording extraction can be seen in Fig. 4.

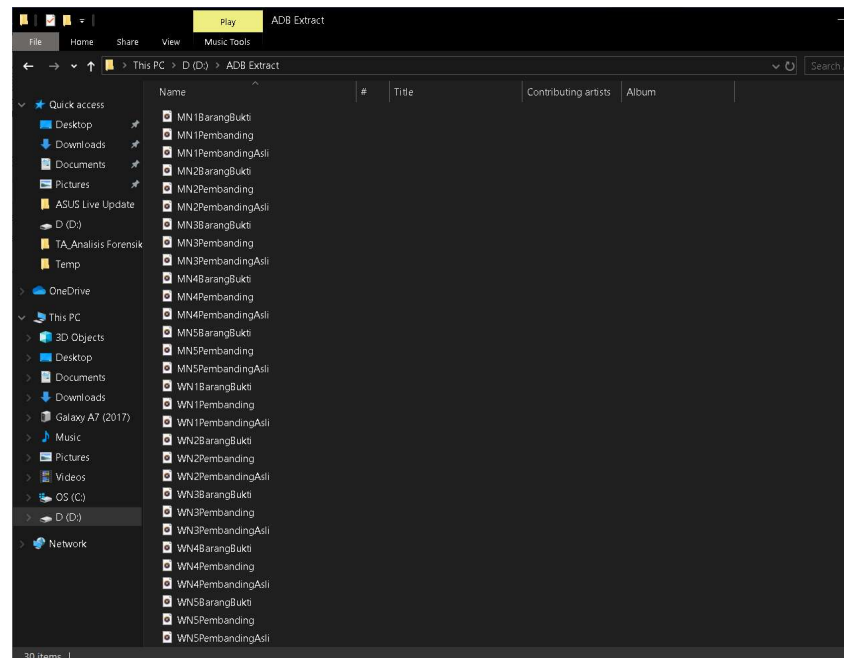Fig. 4. Voice changed recognize voice recording

Fig. 5. The file directory of the extracted data results

After data extraction, the files are stored in one directory file and they are changed to WAV format. Result of all data extraction can be seen in Fig. 5.

B. Audio Forensic Method Analysis

This study uses twenty types of voice recordings. five male voice recordings from Samsung S8 coded as MN 1, MN 2, MN 3, MN 4, MN 5, five male voice recordings from Samsung A7 coded as MN 1', MN 2', MN 3', MN 4', MN 5',five female voice recordings from Samsung S8 coded as WN 1, WN 2, WN 3, WN 4, WN 5, and five female voice recordings from Samsung A7 coded as WN 1', WN 2', WN 3', WN 4', and WN 5'. Each sound recording consists of 20 words according to the standard audio forensic analysis of the Federal Bureau of Investigation (FBI)."*Akan tetapi seperti halnya dengan barang bukti digital lainnya, rekaman suara juga sangat rentan dan mudah untuk dirubah/dimanipulasi baik untuk kepentingan pribadi ataupun kelompok. misalnya menggunakan fasilitas aplikasi perubah suara/voice changer yang banyak tersedia pada google play store*" is an example of a sentence used to test in this study. The sentence is chosen randomly from online articles.

Audacity is used to cut and get the 20 words from the sound record both the voice record of evidence and the recognized voice. Then pitch, formant, spectrogram, ANOVA, and likelihood ratio analysis are conducted to compare each evidence word with each recognized word. Praat is the software tool used for pitch, formant, and spectrogram analysis. Gnumeric spreadsheet is the tool used for the ANOVA method and likelihood ratio. It is important to note that the sound recording to be analyzed has the same sampling rate, so that it does not affect the results of the study [20].

Audio Forensic Method of Male and Female Voice

The Audio Forensic Method of Male and Female Voice result is presented in Table V. Each word of the recognized record is compared with each word of the evidence record for each analysis method (Pitch, ANOVA, Likelihood Ratio, and Spectrogram). Each analysis method will provide the amount of words that are identical and that are not identical. All the identical words from each analysis and are totaled, so are the words that are not identical. If the total of the identical words is higher than the total of the words that are not identical, then the decision is that both voice records are identical. Otherwise, the decision is that both voice records are not identical.

As shown in Table V, all comparative results in the Audio Forensic analysis provide the same decision. Male voice record (MN) and the evidence of male voice record (MN') results after using the analysis are all identical where the female voice record (WN) and evidence of female voice record (WN') are also all identical

The standard operation procedure (SOP) on Audio Forensic Analysis from the Digital Forensic Analyst Team (DFAT) Digital Forensic Laboratory Centre at Bareskrim POLRI is proven to be an effective method for Audio Forensics because it has succeeded in detecting all identical sounds, both men and women. Next, this method will be used for Audio Forensics when voice changer with telephone effects is used as an Anti-Forensic tool.

C. Anti-Forensics Method Analysis

After proving that the Audio Forensic method used has a good performance, now the voice changer with telephone effects will be tried as an Anti-Forensic method. If after the sound that is processed in the voice changer with telephone effects is considered not identical with the Audio Forensic method, then the voice changer with the telephone effect will be considered as an effective Anti-Forensic method.

Anti-Forensic Method of Male and Female Voice

The Anti-Forensic method of male and female voice result is presented in Table VI. Here the recognized voice has been processed with voice changer with telephone effect. The Audio Forensic process is same as the previous test. As in the previous test, the Audio Forensic analysis will have two possible results, which are identical or not identical.

TABLE V
VOICE RECOGNITION RESULT RECOGNIZE VOICE RECORDING (MN/WN)
WITH EVIDENCE VOICE RECORDING (MN'/WN')

| Identity | Class | Tested Words | Analysis | | | | Total | Decision |
			Pitch	ANOVA	Likelihood Ratio	Spectrogram		
MN 1	Identical	20	16	14	14	15	59	Identical
	Not Identical		4	6	6	5	21	
MN 2	Identical	20	14	12	12	13	51	Identical
	Not Identical		6	8	8	7	29	
MN 3	Identical	20	11	10	10	17	48	Identical
	Not Identical		9	10	10	3	32	
MN 4	Identical	20	13	14	14	16	57	Identical
	Not Identical		7	6	6	4	23	
MN 5	Identical	20	11	13	13	15	52	Identical
	Not Identical		9	7	7	5	28	
WN 1	Identical	20	10	11	11	14	46	Identical
	Not Identical		10	9	9	6	34	
WN 2	Identical	20	13	12	12	20	57	Identical
	Not Identical		7	8	8	0	23	
WN 3	Identical	20	10	11	11	12	44	Identical
	Not Identical		10	9	9	8	36	
WN 4	Identical	20	11	12	12	12	47	Identical
	Not Identical		9	8	8	8	33	
WN 5	Identical	20	11	11	11	11	44	Identical
	Not Identical		9	9	9	9	36	

As shown in Table VI, after the recognized voice has been processed with voice changer with telephone effect all comparative results in the Audio Forensic analysis provide different results with the previous test on Table V. The recognized male voice record, and the evidence male voice record are all not identical and the recognized female voice record and the evidence female voice record now are also all not identical.

Based on Anti-Forensic testing with male and female voices, the voices that were previously considered identical by the Audio Forensic method, after going through a voice changer process with telephone effects, are now considered not identical. Therefore, a voice changer with telephone effects can be considered as an effective Anti-Forensic tool.

D. Discussion

Fig. 6 shows a bar chart comparing the results obtained in testing the Audio Forensic method and testing the Anti-Forensic tool. Here three correlations can be made. The first is the correlation between gender and the Audio Forensic method. The second is the correlation between gender and the Anti-Forensic tool. The third is the correlation between the Audio Forensic method and the Anti-Forensic tool.

In correlation between gender and the Audio Forensic method, based on data collected in this research, the Audio Forensic method is more effective with male gender. This is because the average value of the male gender from identical results is higher than the average value of the women gender from identical results, which are 53.4 compared to 47.

TABLE VI
VOICE RECOGNITION RESULT EVIDENCE VOICE RECORDING (MN'/WN') WITH RECOGNIZE VOICE RECORDING
(MN/WN) THAT HAS BEEN CHANGED USING VOICE CHANGER

Identity	Class	Tested Words	Analysis				Total	Decision
			Pitch	ANOVA	Likelihood Ratio	Spectrogram		
MN 1	Identical	20	5	4	4	9	22	Not Identical
	Not Identical		15	16	16	11	58	
MN 2	Identical	20	10	5	5	6	26	Not Identical
	Not Identical		10	15	15	14	54	
MN 3	Identical	20	5	2	2	7	16	Not Identical
	Not Identical		15	18	18	14	64	
MN 4	Identical	20	5	0	0	6	11	Not Identical
	Not Identical		15	20	20	14	69	
MN 5	Identical	20	1	1	1	5	8	Not Identical
	Not Identical		19	19	19	15	72	
WN 1	Identical	20	0	3	3	14	20	Not Identical
	Not Identical		20	17	17	7	60	
WN 2	Identical	20	3	9	9	13	34	Not Identical
	Not Identical		17	11	11	7	46	
WN 3	Identical	20	5	1	1	5	12	Not Identical
	Not Identical		15	19	19	15	68	
WN 4	Identical	20	3	1	1	4	9	Not Identical
	Not Identical		17	19	19	16	71	
WN 5	Identical	20	6	3	3	6	18	Not Identical
	Not Identical		14	18	18	14	62	

In correlation between gender and the Anti-Forensic tool, the average value of the male gender from not identical results is higher than the average value of the women gender from not identical results, which are 63.4 compared to 61.4. However, through a t-test, the t stat value of the two ranges of values is not great enough to conclude that the two averages are different. Hence, for the Anti-Forensic tool, it cannot be concluded that the Anti-Forensic tool is more effective on male gender.

Additionally, system performance is dependent on the voice of the person being tested, because several people have different variations of the voice [21].

V. CONCLUSION

Voice Changer with telephone effect is an effective Anti-Forensics tool. This is proven after the voice changer with telephone effect can change the results that were previously identical with the proven Audio Forensic method. Additionally, based on the data collected in this research, the Audio Forensic method is more effective on identifying male voices compared to female voices, where the average score is 53.4 and 47 respectively. However, the Anti-Forensic tool is not proven to be more effective on any gender because the results of the t-Test did not reject the null hypothesis. For future work, the influence of the alpha (α) and beta (β) scale factors provided by the voice changer in the Anti-Forensic tool's effectiveness can be measured
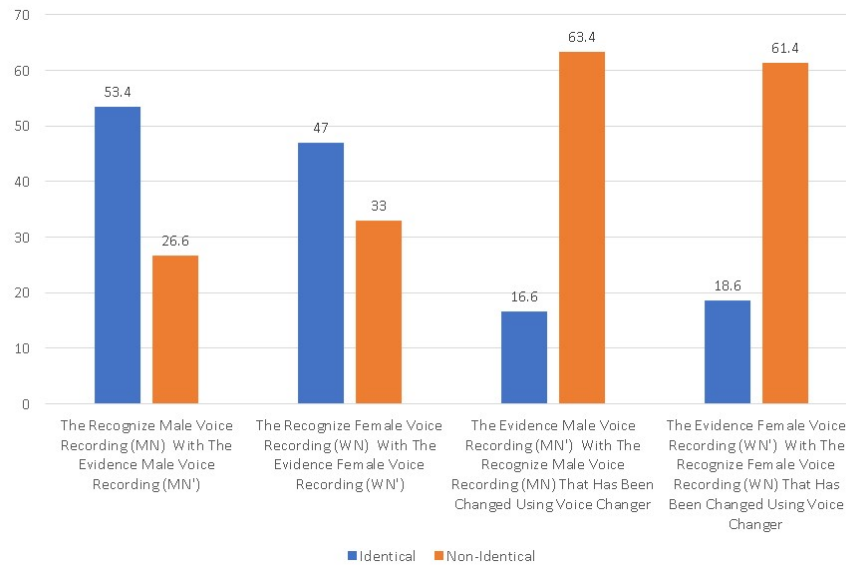
Fig. 6. Comparison of the average value of Audio Forensics without Anti-Forensics and with Anti-Forensics.

REFERENCES

[1] K. Conlan, I. Baggili, and F. Breitinger, "Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy," *DFRWS 2016 USA - Proc. 16th Annu. USA Digit. Forensics Res. Conf.*, vol. 18, no. December 2015, pp. S66–S75, 2016, doi: 10.1016/j.diin.2016.04.006.

[2] J. Sarwono, M. I. Mandasari, and Suprijanto, "Forensic Speaker Identification: An experience in Indonesians court," *20th Int. Congr. Acoust. 2010, ICA 2010 - Inc. Proc. 2010 Annu. Conf. Aust. Acoust. Soc.*, vol. 5, no. June 2015, pp. 3861–3863, 2010.

[3] S. Jadhav, R. Patole, and P. Rege, "Audio Splicing Detection using Convolutional Neural Network," *2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019*, pp. 1–5, 2019, doi: 10.1109/ICCCNT45670.2019.8944345.

[4] H. Zhao, Y. Chen, R. Wang, and H. Malik, "Anti-Forensics of Environmental-Signature-Based Audio Splicing Detection and Its Countermeasure via Rich-Features Classification," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 7, pp. 1603–1617, 2016, doi: 10.1109/TIFS.2016.2543205.

[5] R. K. Chan, "Speaker discrimination: Citation tones vs. coarticulated tones," *Speech Commun.*, vol. 117, no. June 2019, pp. 38–50, 2020, doi: 10.1016/j.specom.2019.06.006.

[6] P. R. Bevinamarad and M. S. Shirldonkar, "Audio Forgery Detection Techniques: Present and Past Review," no. Icoei, pp. 613–618, 2020, doi: 10.1109/icoei48184.2020.9143014.

[7] A. Subki, B. Sugiantoro, and Y. Prayudi, "Analisis Rekaman Suara Voice Changer dan Rekaman Suara Asli Menggunakan Metode Audio Forensik," *Indones. J. Netw. Secur.*, vol. 7, no. 1, 2018.

[8] W. Zhenhua, W. Yingli, and D. S. Nanda, "Forensic Linguistic Inquiry into the Validity of F0 as Discriminatory Potential in the System of Forensic Speaker Verification," *J. Forensic Sci. Crim. Investig.*, vol. 5, no. 3, pp. 1–8, 2017, doi: 10.19080/jfsci.2017.05.555664.

[9] M. A. Wani, A. AlZahrani, and W. A. Bhat, "File system anti-forensics – types, techniques and tools,"

Comput. Fraud Secur., vol. 2020, no. 3, pp. 14–19, 2020, doi: 10.1016/S1361-3723(20)30030-0.

[10]   F. Sulianta, *Komputer Forensik*. Indonesia: Elex media Komputindo, 2013.

[11]   M. A. Qureshi and E. S. M. El-Alfy, "Bibliography of digital image anti-forensics and anti-anti-forensics techniques," *IET Image Process.*, vol. 13, no. 11, pp. 1811–1823, 2019, doi: 10.1049/iet-ipr.2018.6587.

[12]   P.-J. Moon, "On the Availability of Anti-Forensic Tools for Android Smartphones," *J. Korea Inst. Electron. Commun. Sci.*, vol. 8, no. 6, pp. 855–861, 2013, doi: 10.13067/jkiecs.2013.8.6.855.

[13]   Q. Yan, R. Yang, and J. Huang, "Robust copy-move detection of speech recording using similarities of pitch and formant," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 9, pp. 2331–2341, 2019, doi: 10.1109/TIFS.2019.2895965.

[14]   A. Subki, B. Sugiantoro, and Y. Prayudi, "Membandingkan Tingkat Kemiripan Rekaman Voice Changer Menggunakan Analisis Pitch, Formant dan Spectogram," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 1, p. 17, 2018, doi: 10.25126/jtiik.201851500.

[15]   V. R. . Putri, "Analisis Rekaman Suara Menggunakan Teknik Audio Forensik Untuk Keperluan Barang Bukti Digital," *Unnes Phys. J.*, vol. 3, no. 1, 2014.

[16]   H. Cao, "the Role of the First Five Formants in Three Vowels of," no. August, pp. 617–621, 2019.

[17]   M. N. Al-Azhar, *Digital Forensic Practical Guidelines For Computer Investigation*. Jakarta, 2010.

[18]   Y. Ren, D. Liu, Q. Xiong, J. Fu, and L. Wang, "Spec-ResNet: A General Audio Steganalysis scheme based on Deep Residual Network of Spectrogram," pp. 1–12, 2019.

[19]   R. C. Maher, "Audio Forensic Examination: Authenticity, enhancement, and interpretation," *Ieee Signal Process. Mag. [84]*, no. March, pp. 84–94, 2009.

[20]   R. R. Huizen, N. K. D. A. Jayanti, and D. P. Hostiadi, "Analisis Pengaruh Sampling Rate Dalam Melakukan Identifikasi Pembicara Pada Rekaman Audio," *Konf. Nas. Sist. Inform.*, 2015.

[21]   V. Hughes, P. Harrison, P. Foulkes, P. French, C. Kavanagh, and E. S. Segundo, "The individual and the system: Assessing the stability of the output of a semiautomatic forensic voice comparison system," 2018, doi: 10.21437/Interspeech.2018-1649.