

Implementasi *Voice Recognition* pada Sistem Pengawasan Anak-Anak Dalam Berkata Kasar Melalui *Smartphone* Dengan Koneksi WiFi

Yunita Retno Dewi¹, Ferry Satria², Maya Rahayu³

¹Jurusan Teknik Elektro, Politeknik Negeri Bandung, Bandung 40012
E-mail : yunita.retno.tcom17@polban.ac.id

²Jurusan Teknik Elektro, Politeknik Negeri Bandung, Bandung 40012
E-mail : satriapolban@gmail.com

³Jurusan Teknik Elektro, Politeknik Negeri Bandung, Bandung 40012
E-mail : mayarahayu@polban.ac.id

ABSTRAK

Kata-kata kasar adalah kata-kata yang umumnya bersifat tidak sopan, kasar atau menyinggung. Pada zaman sekarang, seseorang yang berkata kasar sudah tidak asing didengar. Kini anak-anak pun sudah mulai berbicara kasar dan hal tersebut sangat fatal jika tidak dididik atau dicegah dengan baik. Seiring dengan kemajuan teknologi yang semakin canggih, maka dapat dimanfaatkan dalam membuat sistem yang bertujuan untuk mempermudah orangtua dalam mengawasi anak-anak dalam berkata kasar, dapat menimbulkan rasa malu dan meminimalisir seseorang dalam berkata kasar. Sistem ini bekerja saat *Voice Recognition V3* dapat mengenali suara jika terdapat seseorang yang berkata kasar. *Arduino UNO* digunakan sebagai mikrokontroler yang menggunakan bahasa pemrograman C dihubungkan dengan *Voice Recognition V3* dan *Buzzer* sebagai *output* jika seseorang berkata kasar. Tidak hanya itu, alat ini menggunakan *blynk app* untuk mengirimkan informasi berupa notifikasi *e-mail* pada *smartphone* orangtua dengan mengakses *IP Address ESP32-CAM* yang digunakan untuk menampilkan keadaan suatu ruangan. *ESP32-CAM* merupakan mikrokontroler yang sudah memiliki modul *wifi* dan terdapat *OV2640DS* yang digunakan untuk pengawasan keadaan suatu ruangan jika terdapat seseorang berkata kasar. Berdasarkan dari hasil pengujian yang dilakukan jika dalam keadaan ideal, *Voice Recognition V3* dapat mengenali suara sebesar 94,29%. Sedangkan saat terdapat *noise*, *Voice Recognition V3* dapat mengenali suara sebesar 25,7%. Dan jika dilakukan pengujian pengenalan suara dengan orang yang berbeda memiliki tingkat keberhasilan sebesar 31,43%. Dapat disimpulkan bahwa *Voice Recognition V3* dapat bekerja secara ideal jika keadaan suatu ruangan yang tidak banyak *noise*. Dan jika *Voice Recognition V3* dapat bekerja dengan baik, maka *buzzer* akan berbunyi dan mengirimkan notifikasi ke *e-mail* orangtua.

Kata Kunci

Voice Recognition V3, Berkata kasar, Arduino UNO, Buzzer, E-mail, dan ESP32-CAM

1. PENDAHULUAN

Kata-kata kasar adalah kata-kata yang secara budaya bersifat tidak sopan, kasar atau menyinggung. Umumnya kata-kata kasar berkaitan dengan penghinaan terhadap orang lain dan berkaitan dengan perasaan yang kuat terhadap sesuatu. Permasalahan yang sering dialami adalah banyak sekali masyarakat khususnya anak-anak yang berkata kasar dengan meluapkan rasa kesal kepada seseorang atau hal yang tidak enak didengar. Kebiasaan berbicara kasar erat hubungannya dengan lingkungan sekitarnya. Apabila seseorang berkata kasar di depan anak-anak bisa membahayakan dikarenakan anak-anak menganggap bahwa semua yang dilakukan orang dewasa bisa dijadikan sebagai hal yang wajar. Oleh karena itu, pengawasan orangtua sangat penting karena karakter anak tumbuh dari lingkungan terdekatnya seperti lingkungan keluarga.

Jika tidak dididik dengan baik akan menumbuhkan karakter anak yang kurang baik. Untuk mengatasi masalah tersebut dapat memanfaatkan teknologi pengenalan suara yang ada untuk mengawasi anak-anak dalam berkata kasar agar orangtua dapat mengawasi keadaan suatu ruangan.

2. PUSTAKA TERKAIT

Pada permasalahan ini dari kata-kata kasar tersebut merupakan hasil peniruan anak-anak dari apa yang didengarnya. Pada usia awal anak mulai belajar membaca, banyak menyerap kata melalui mendengarkan dan meniru kata-kata yang diucapkan oleh lingkungannya.

Terdapat beberapa sistem yang pernah dibuat untuk memecahkan masalah ini yaitu dengan pemanfaatan *text mining* pada studi kasus Radio Pikiran Rakyat

(PR) 107.5 FM Bandung yang berfungsi untuk mendeteksi kata-kata kotor pada SMS yang berasal dari pendengar lalu menentukan layak atau tidak layak makna dari kata tersebut untuk dibacakan *on air*. Dan menerapkan *Naïve Bayes Classifier* untuk mengoptimalkan proses pendeteksian kata tidak layak [1]. Sistem selanjutnya yaitu dengan melakukan perancangan sistem monitoring dan manajemen rekaman sehingga didapat sistem yang murah. Cara kerja sistem monitoring dengan menangkap sinyal radio FM, kemudian direkam. Data yang telah terekam disimpan, kemudian dilakukan pengarsipan. Sistem ini bertujuan untuk meminimalisir pelanggaran pada siaran radio di Indonesia dengan membuat sistem monitoring mandiri yang dapat membantu proses monitoring yang dilakukan pemerintah [2].

Mengeluarkan bau yang tidak sedap dari suatu USB yang telah dihubungkan ke komputer jika seseorang mengetikkan kata-kata kasar. Hal tersebut dilakukan untuk pencegahan dalam melakukan pencarian atau pengetikan kata di dunia maya [3].

Akhmad Wahyu Dani dkk. dengan judul "Perancangan Aplikasi Voice Command Recognition Berbasis Android dan Arduino UNO". Alat ini memiliki kelebihan dan kekurangannya yaitu: Kelebihannya adalah mudah digunakan, penerapannya sederhana. Sedangkan kekurangannya adalah sangat bergantung kepada server google untuk dapat mengartikan perintah suara yang ada [4].

Jarak keberhasilan pemberian perintah dipengaruhi oleh jarak pemberian *sample* suara. Pemberian perintah pada jarak ≥ 1 meter akan mengalami kesulitan pendeteksian perintah. Dan ketika pemberian perintah dengan *noise* ≥ 47 dB, tingkat keberhasilan mengalami penurunan signifikan. Itu disebabkan perintah suara yang diterima bercampur dengan *noise* dan mengakibatkan kesulitan untuk melakukan pengenalan perintah. Penggunaan *microphone* yang berbeda dapat mempengaruhi keakuratan pengenalan suara, itu dikarenakan sensitivitas tiap *microphone* berbeda [4].

Berdasarkan dengan karya/sistem terdahulu dan pustaka terkait, dikembangkanlah suatu karya/sistem yaitu "Implementasi Voice Recognition pada Sistem Pengawasan Anak-anak dalam Berkata Kasar melalui Smartphone dengan Koneksi Wifi" agar dapat meningkatkan kesadaran anak-anak bahwa berbicara kata-kata kasar itu sangatlah fatal yang berdampak di lingkungan sekitar. Alat ini diharapkan dapat menimbulkan rasa malu dan meminimalisir seseorang yang berkata kasar karena *alarm* tersebut berbunyi.

3. METODOLOGI PELAKSANAAN

Pada tahap persiapan realisasi perangkat keras diperlukan tahap-tahap yang dilakukan dalam membuat *prototype* tersebut.

3.1 Ilustrasi Gambar



Gambar 1. Ilustrasi Gambar

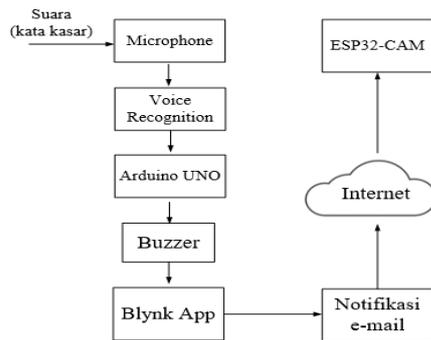
Gambar 1 merupakan ilustrasi gambar dari sistem yang diusulkan. Yaitu, suara (kata kasar) diterima oleh *microphone* akan diteruskan ke *Voice Recognition V3* untuk mengenali suara(kata kasar) yang telah dihubungkan dengan *Arduino UNO* untuk memberikan *output* berupa *buzzer*. *Buzzer* tersebut digunakan agar orangtua mengetahui bahwa terdapat anak-anak yang telah berkata kasar. Saat *buzzer* menyala, orangtua mendapatkan notifikasi *e-mail* yang terhubung dengan *blynk app*. Notifikasi *e-mail* tersebut berisi peringatan kepada orangtua untuk melakukan pengawasan anak-anak dalam berkata kasar dengan mengakses *IP Address* pada *ESP32-CAM*. *ESP32-CAM* harus terhubung terlebih dahulu dengan *wifi* agar dapat mengawasi keadaan suatu ruangan, contohnya seperti kamar tidur, ruang tamu, dll.

3.2 Algoritma yang Digunakan

Algoritma yang digunakan dalam pembuatan *prototype* sebagai berikut:

1. Melakukan *train* terlebih dahulu kata-kata kasar yang akan digunakan pada *Voice Recognition V3*.
2. Kata kasar yang telah di-*train* akan diolah pada *Arduino UNO* yang sudah dihubungkan dengan *Voice Recognition V3* dan menghasilkan *output* berupa *buzzer*.
3. Setelah itu, untuk menghubungkan komponen yang terhubung dengan *Arduino UNO* dengan *ESP32-CAM* dibuatlah aplikasi menggunakan *blynk app* dengan koneksi *port USB*. Untuk membuat notifikasi *e-mail* pada *blynk app* ini dengan menambahkan *widget e-mail* yang sudah tersedia pada *blynk app* agar dapat terhubung dengan *e-mail*. Notifikasi *e-mail* berisi peringatan kepada orangtua bahwa terdapat seseorang berkata kasar dan *link IP Address ESP32-CAM* yang dapat diakses.
4. *ESP32-CAM* harus terhubung terlebih dahulu dengan *wifi* agar dapat mengawasi keadaan suatu ruangan, contohnya seperti kamar tidur, ruang tamu, dll.

3.3 Blok Diagram yang Digunakan



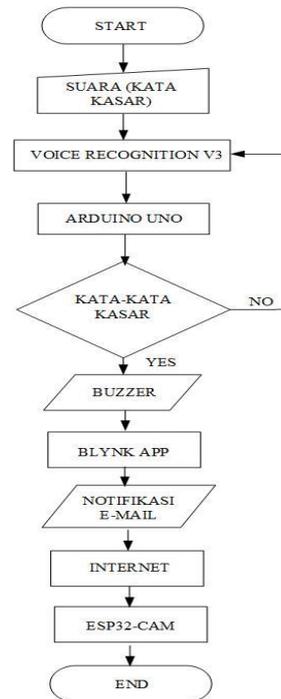
Gambar 2. Blok Diagram yang Digunakan

Gambar 2 adalah blok diagram yang digunakan secara keseluruhan. Kata-kata kasar yang digunakan pada *prototype* ini adalah anjing, goblok, tolol, bego, bangsat, kampret dan brengsek. Kata-kata tersebut dikategorikan sebagai kata-kata kasar yang sering didengar dan dapat membuat seseorang merasa direndahkan, disakiti, dihina, dll. Contohnya seperti kata “Anjing” memiliki makna yang lain, bukan sebagai nama hewan melainkan jadi kata makian serta dapat menggambarkan kekesalan. Sehingga para penutur bahasa melakukan perbandingan dengan binatang yang direndahkan[6]. Suara (kata kasar) akan dikenal oleh *Voice Recognition V3* yang diteruskan ke *Arduino UNO*. Jika *Voice Recognition V3* mengenali suara (kata kasar), *Arduino UNO* akan mengolah data yang akan dikirim ke *buzzer* sebagai *output*. Dari *buzzer* akan diteruskan ke *blynk app* yang digunakan untuk menghubungkan antara komponen yang telah terhubung dengan *Arduino UNO* dan notifikasi *e-mail*. Saat *buzzer* berbunyi, akan mengirimkan notifikasi *e-mail* berupa peringatan bahwa terdapat anak-anak yang berkata kasar dan dapat mengakses *IP Address* dari *ESP32-CAM* yang sudah terhubung dengan *wifi* untuk melakukan pengawasan keadaan suatu ruangan.

3.4 Perancangan Diagram Alir Keseluruhan

Gambar 3 merupakan perancangan diagram alir sistem secara keseluruhan. Diagram alir ini akan direalisasikan pada mikrokontroler. Suara kata-kata kasar harus di-*train* terlebih dahulu pada *Voice Recognition V3* agar dapat mengenali kata-kata kasar. Jika *Voice Recognition V3* mengenali suara (kata kasar) akan diteruskan ke *Arduino UNO* untuk menghasilkan *output* berupa *buzzer*. Saat *buzzer* berbunyi, orangtua akan mendapatkan notifikasi *e-mail* pada *smartphone* yang menunjukkan terdapat anak-anak yang berkata kasar. Dan orangtua dapat melakukan *live streaming* keadaan suatu ruangan pada *smartphone* yang telah terhubung dengan *wifi* dan *IP Address Camera* dari komponen *ESP32-CAM*. Untuk mendapatkan notifikasi *e-mail* tersebut, *blynk app* digunakan untuk

menghubungkan antara komponen *Arduino UNO* dengan notifikasi *e-mail*.



Gambar 3. Perancangan Diagram Alir Keseluruhan

3.5 Realisasi

Pada tahap realisasi ini akan menjelaskan bagaimana perealisasi *prototype* dari perealisasi perangkat lunak sampai menjadi kemasan alat.

3.5.1 Realisasi Perangkat Keras

3.5.1.1 Realisasi PCB



Gambar 4. Realisasi PCB

Gambar 4 merupakan realisasi *PCB* dengan menggunakan *PCB matriks* yang berukuran 7,5 cm x 5 cm. Pada *PCB matriks* hanya digunakan oleh komponen *Voice Recognition V3* dengan *buzzer*.

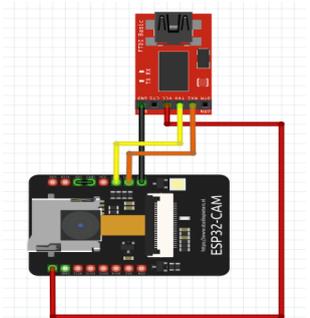
3.5.1.2 Perakitan



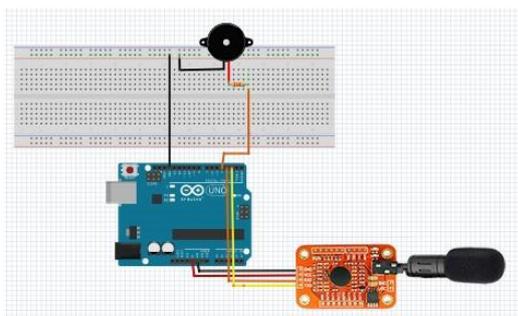
Gambar 5. Perakitan

Gambar 5 merupakan perakitan dari beberapa komponen yaitu *Arduino UNO* yang dihubungkan *buzzer* yang digunakan sebagai *output* dari *Voice Recognition V3*. Untuk *board* yang lainnya yaitu komponen *FTDI* dengan *ESP32CAM*.

3.5.1.3 Pengkabelan



Gambar 6. Pengkabelan *ESP32CAM* dan *FTDI*



Gambar 7. Pengkabelan *Arduino UNO* dan *Voice Recognition V3*

3.5.1.4 Realisasi Kemasan Alat

Gambar 8 merupakan realisasi perangkat keras dengan *box* yang berukuran 18 cm x 11 cm x 6 cm. Pada tahap realisasi kemasan alat menggunakan *box* yang telah dilubangi untuk *USB Port*, *camera ESP32-CAM* dan *microphone*.



Gambar 8. Realisasi Kemasan Alat

4. HASIL PENGUJIAN

Pengujian ini dilakukan untuk mengetahui seberapa sensitif *Voice Recognition V3* dalam mendeteksi kata-kata kasar jika dilakukan saat keadaan suatu ruangan tidak terdapat *noise* dan terdapat *noise* dengan jarak 5-150 cm. Selain itu, pengujian dilakukan dengan mengenali kata kasar yang dilakukan oleh 5 orang yang berbeda dengan jarak sejauh 25 cm.

4.1 Gambaran Pelaksanaan Pengujian

Cara melaksanakan pengujiannya adalah sebagai berikut:

1. *Prototype* diletakkan di tempat yang ingin diawasi oleh orangtua.
2. Jika seseorang berkata kasar sesuai dengan kata kasar yang telah di-*train* pada *Voice Recognition V3* maka *buzzer* akan berbunyi dan orangtua mendapatkan notifikasi pada aplikasi *smartphone* yang telah di-*install*.
3. Untuk pengujian dilakukan berdasarkan jarak. Dilakukan pengujian berdasarkan jarak antara seseorang yang berkata kasar dengan *prototype* yang telah diletakkan di suatu ruangan. Pengujian yang dilakukan dengan jarak sejauh 5 cm, 10 cm, 50 cm, 100 dan 150 cm dalam keadaan tidak terdapat *noise* dan terdapat *noise*.
4. Pengujian selanjutnya dilakukan pengenalan suara kata kasar oleh 5 orang yang berbeda dengan jarak sejauh 25 cm.

Berikut adalah hasil pengujian jarak ideal dalam mendeteksi kata kasar.

Tabel 1. Hasil Pengujian Jarak Ideal dalam Mendeteksi Kata Kasar

Kata Perintah	5 cm	10 cm	50 cm	100 cm	150 cm	Kata Kasar yang Dikenali (%)
Anjing	✓	✓	✓	✓	✓	100
Goblok	✓	✓	✓	✓	✓	100
Tolol	✓	✓	✓	✓	-	80
Bego	✓	✓	✓	✓	✓	100
Bangsas	✓	✓	✓	✓	✓	100
Kampret	✓	✓	✓	✓	✓	100
Brengsek	✓	✓	✓	✓	-	80
Rata-rata keberhasilan kata-kata kasar yang dikenali (%)						94,29

Berdasarkan hasil pengujian pada tabel 1 menunjukkan bahwa rata-rata tingkat keberhasilan kata-kata kasar yang dapat dikenali sebesar 94,29 %. Jadi, dapat disimpulkan bahwa semakin jauh jarak antara seseorang yang berkata kasar dengan *Voice Recognition V3*, maka kesensitifitasannya akan berkurang. Jika lebih dari 100 cm maka *Voice Recognition V3* tidak mengenali suara kata-kata kasar dengan baik dikarenakan kesensitifitasan yang semakin menurun.

Tabel 2. Hasil Pengujian Jarak Ideal dalam Mendeteksi Kata Kasar (Terdapat Noise)

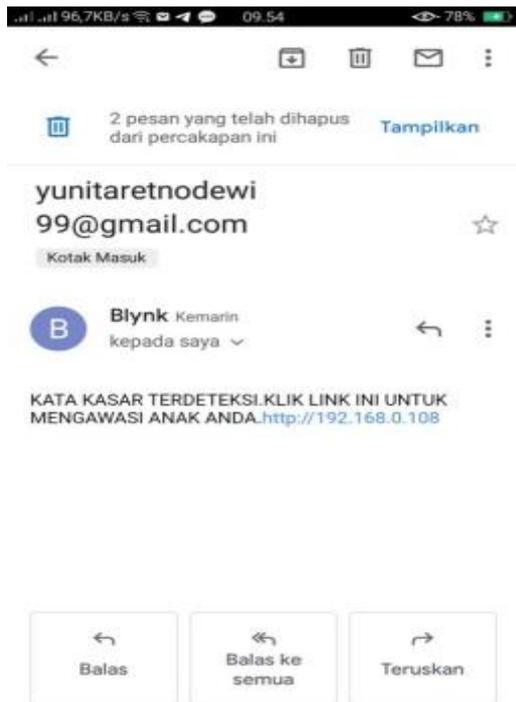
Kata Perintah	5 cm	10 cm	50 cm	100 cm	150 cm	Kata Kasar yang Dikenali (%)
Anjing	✓	✓	-	-	-	40
Goblok	✓	-	-	-	-	20
Tolol	✓	-	-	-	-	20
Bego	✓	✓	-	-	-	40
Bangsas	✓	-	-	-	-	20
Kampret	✓	-	-	-	-	20
Brengsek	✓	-	-	-	-	20
Rata-rata keberhasilan kata-kata kasar yang dikenali (%)						25,7

Berdasarkan hasil pengujian pada tabel 2 menunjukkan bahwa rata-rata tingkat keberhasilan kata-kata kasar yang dapat dikenali sebesar 25,7%. Jadi, dapat disimpulkan bahwa jika keadaan suatu ruangan terdapat *noise* maka *Voice Recognition V3* tidak bekerja secara ideal. Dikarenakan *Voice Recognition V3* atau *microphone* yang digunakan tidak memiliki kesensitifitasan yang tinggi.

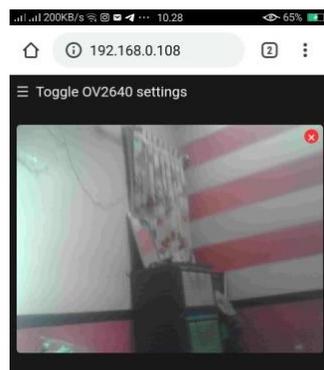
Tabel 3. Hasil Pengujian kata kasar dengan suara berbeda

Kata Perintah	Suara					Kata Kasar yang Dikenali (%)
	1	2	3	4	5	
Anjing	✓	✓	✓	-	-	60
Goblok	✓	-	-	-	-	20
Tolol	✓	-	-	-	-	20
Bego	✓	-	-	-	✓	40
Bangsas	✓	✓	-	-	-	40
Kampret	✓	-	-	-	-	20
Brengsek	✓	-	-	-	-	20
Rata-rata keberhasilan kata-kata kasar yang dikenali (%)						31,43

Keterangan: Suara 1 dilakukan oleh seseorang yang melakukan *train* terlebih dahulu pada *Voice Recognition V3*. Berdasarkan hasil pengujian pada tabel 3 menunjukkan bahwa rata-rata tingkat keberhasilan kata-kata kasar yang dapat dikenali sebesar 31,43 %. Jadi, dapat disimpulkan bahwa *Voice Recognition V3* dapat mengenali kata-kata kasar yang telah di-*train* terlebih dahulu. Kecuali, jika suara saat di-*train* sama dengan atau mirip dengan suara yang dilakukan dalam pengujian.



Gambar 9. Notifikasi E-mail pada Smartphone



Gambar 10. Live Streaming ESP32-CAM

Pada gambar 9 merupakan notifikasi *e-mail* pada *smartphone* orangtua jika terdapat seseorang yang telah berkata kasar. Notifikasi *e-mail* memberikan peringatan kepada orangtua untuk mengawasi anaknya dengan memberikan *link IP Address ESP32-CAM* untuk melakukan *live streaming* keadaan suatu ruangan yang terdapat pada gambar 10.

5. KESIMPULAN DAN SARAN

Berdasarkan dari hasil pengujian yang telah dilakukan dapat disimpulkan bahwa :

1. Berdasarkan hasil pengujian, *Voice Recognition V3* dapat bekerja sebesar 94,29 % dalam keadaan ideal dengan jarak dari 5 cm sampai dengan 150 cm. Sedangkan, jika dalam keadaan terdapat *noise Voice Recognition V3* hanya dapat bekerja dengan tingkat keberhasilan sebesar 25,7 %.

2. Pengujian *Voice Recognition V3* yang dilakukan dengan suara 5 orang yang berbeda memiliki tingkat keberhasilan hanya 31,43 %. Dikarenakan hasil pengujian menunjukkan bahwa suara 5 orang yang berbeda tersebut tidak sama dengan suara saat di-*train* pada *Voice Recognition V3*.
3. Notifikasi *e-mail* dapat bekerja dengan baik, jika *Voice Recognition V3* dapat mendeteksi kata kasar saat *buzzer* berbunyi atau terdapat seseorang yang berkata kasar.

Berdasarkan hasil dari pengujian dan kesimpulan yang didapat, untuk mendapatkan hasil yang lebih baik maka penulis memberikan saran sebagai berikut:

1. Dengan menambahkan hasil *capture* suatu ruangan tersebut dan dikirimkan ke *e-mail* orangtua jika terdapat seseorang yang berkata kasar.
2. Dengan meningkatkan kesensitifan *Voice Recognition* atau dengan mengganti *microphone* yang lebih sensitif.
3. Untuk menghasilkan kualitas gambar yang lebih baik saat *live streaming* keadaan suatu ruangan, maka gunakan modul kamera yang memiliki kualitas yang lebih tinggi dari.

DAFTAR PUSTAKA

- [1] Zulfikar, Wildan Budiawar, "Pemanfaatan Text Mining Pada Berita Dengan Metode Naive Bayes Classifier untuk Deteksi Kata Tidak Layak: Studi Kasus Radio PR 107.5 FM Bandung," 2011.
- [2] Safar, Rusli Nasrullah, "Perancangan Sistem Akuisisi dan Manajemen Data," 2017.
- [3] DetikInet. (2010, Juni) [Online]. Available: <https://inet.detik.com/consumer/d-1384392/cegah-kalimat-kotor-komputer-keluakan-bau-tak-sedap> [Diakses 19 Desember 2019].
- [4] Dani, Akhmad Wahyu, "Perancangan Aplikasi Voice Command Recognition Berbasis Android dan Arduino UNO," *Jurnal Teknologi Elektro Universitas Mercu Buana.*, vol. 7, no.1, 2016.
- [5] Jawaami, Arfian Jamul. (2017, Oktober) [Online]. <https://ayobandung.com/read/2017/10/14/24694/mengapa-anjing-jadi-kata-makian> [Diakses 11 Februari 2020].