



PROPOSAL PENELITIAN

KUALIFIKASI

**KLASIFIKASI GENRE FILM INDONESIA BERDASARKAN CUPLIKAN
MENGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)**

META MEYSAWATI

99216026

PROGRAM DOKTOR TEKNOLOGI INFORMASI

UNIVERSITAS GUNADARMA

2022

DAFTAR ISI

Cover	i
Daftar Isi	ii
BAB I. PENDAHULUAN	
1.1. Latar Belakang Masalah	1
1.2. Rumusan dan Batasan Masalah Penelitian	3
1.3. Tujuan Penelitian	3
1.4. Kontribusi Penelitian	3
Bab 2. TINJAUAN PUSTAKA	
2. 1 Tren Film Di Indonesia	5
2. 2 Pengenalan Video dan Audio	6
2. 3 Klasifikasi	7
2.3.1 Metode <i>Convolutional Neural Network</i> (CNN)	7
2.3.2 Metode <i>Support Vector Machine</i> (SVM)	8
2.4 Perkembangan Penelitian Terkait	
2.4.1 Implementasi Metode <i>Convolutional Neural Network</i> (CNN) Pada Klasifikasi Video	8
2.4.2 Implementasi Metode <i>Support Vector Machine</i> (SVM) Pada Klasifikasi Video	10
2.5 Perbandingan Penelitian	10
Bab 3. METODOLOGI PENELITIAN	
3.1 Tahapan Penelitian	14
3.2 Rencana Kerja	17
DAFTAR PUSTAKA	

1. PENDAHULUAN

1.1 Latar Belakang Masalah

Penggunaan suara dalam *data science* telah banyak dilakukan pada beberapa penelitian untuk mendeteksi jenis objek yang diteliti yang kemudian dilakukan pengelompokan atau pengkategorian berdasarkan dari jenis objek tersebut. Suara dapat diperoleh dari media video, ini didasarkan kepada pengertian Video yang merupakan media elektronik yang mampu menggabungkan teknologi audio dan visual secara bersama sehingga menghasilkan suatu tayangan yang dinamis dan menarik. Salah satu unsur yang terdapat pada video diantaranya teks, gambar, dan suara. Salah satu manfaat video adalah sebagai media promosi, karena Promosi merupakan salah satu penentu keberhasilan suatu program pemasaran pada sektor perfilman (Kominfo, 2021).

Dalam hal ini, peneliti akan melakukan pengkategorian menggunakan suara untuk melakukan pengklasifikasian film berdasarkan dari cuplikan film. Penelitian mengenai hal ini, banyak ditemukan apabila dalam pengkategorian film dilakukan melalui ulasan film. Namun tidak sedikit dari peneliti menggunakan data suara sebagai pengolah data untuk pengklasifikasian. Klasifikasi merupakan cara mengelompokkan data sesuai dengan ciri-ciri atau karakteristik data tersebut (Aji Prasetya Wibawa, 2018). Berbagai macam data yang diolah dengan metode yang efektif dan efisien akan menghasilkan informasi yang akurat. Sejalan dengan perkembangan teknologi, bermacam metode pengolahan data juga dikembangkan. Hal ini untuk memenuhi kebutuhan akan metode yang digunakan sesuai dengan tipe data yang akan diolah.

Film merupakan bentuk dominan dari komunikasi massa visual dalam hal memberikan pesan tersirat yang ingin disampaikan oleh pencipta film lewat audio visual (Muslimah, N, dkk, 2019). Film juga dapat menjadi sebuah representasi dari masyarakat, dimana dalam setiap isi cerita film tersebut pasti mengambil dari suatu kebudayaan yang terdapat pada lingkungannya. Film seringkali dilihat sebagai salah satu bentuk hiburan yang menyediakan beragam jenis pilihan atau disebut dengan “genre”. Genre merupakan kategori dari bentuk seni yang digunakan untuk menyebutkan subset dari film (Shafirra, N. A, dkk, 2020). Selain itu, film juga sebagai sarana informasi seperti peristiwa dan kondisi masyarakat dari berbagai macam dunia, salah satunya perfilman di Indonesia.

Dalam lima tahun terakhir, film Indonesia semakin beragam dengan menghadirkan aktualitas dan konteks melalui tema dan genre. Film yang beredar tetap didominasi oleh genre umum yang sudah berlangsung lama, yaitu drama, horor, dan komedi serta penggabungan antara ketiganya (Film Indonesia, 2010). Namun berdasarkan sumber informasi dari film Indonesia terdapat 14 genre film yaitu laga, drama, horor, perang, anak-anak, fantasi, legenda, petualangan, komedi, mistik, remaja, musikal, thriller dan animasi. Dari banyaknya genre yang disediakan oleh film maka tidak sedikit dari penikmat film masih bingung dan sulit untuk membedakan genre-genre film tersebut.

Tidak sedikit upaya yang dilakukan untuk menayangkan sebuah film salah satunya adalah dengan upaya promosi film yang akan segera ditayangkan berupa video yang memuat plot dari inti cerita dalam film, termasuk pemeran, sutradara, produser distributor dan waktu tayang film tersebut. Yang mana cuplikan ini sudah mengandung beberapa unsur penting dalam film (Film Indonesia, 2010). Cuplikan film disajikan dalam bentuk video dengan durasi 2-3 menit dan sudah menjelaskan inti cerita pada film tersebut. Hal ini yang membuat para peminat film untuk menonton cuplikan terlebih dahulu sebelum menentukan film apa yang hendak ditonton. Selain itu pun, penayangan cuplikan akan lebih efektif dalam hal mempromosikan sebuah film (Widya, D. A, 2017). Sehingga tidak sedikit penikmat film memilih untuk menonton cuplikan karena penyampaian pesan disajikan dalam bentuk media audio visual yaitu berupa video yang membuat penasaran pada film tersebut (IDN Times, 2019). Konten audio dalam trailer film berisi banyak informasi berguna untuk memprediksi genre film, mulai dari suara latar belakang dan latar depan seperti suara lingkungan kendaraan, hewan, alarm, perangkat dll yang bersamaan dengan dialog dan musik. Sehingga konten audio menyediakan banyak fitur yang menonjol untuk memprediksi genre film. Fitur audio ini memberikan pemetaan yang berguna dengan genre film. Misalnya, suara seperti teriakan, tangisan, tawa setan, dll. peta dengan genre 'horor'. Suara ledakan, kebisingan kendaraan, dll., petakan dengan ini genre 'aksi' (Sharma, A, dkk, 2021).

Peneliti yang dilakukan (Sharma, A., Jindal, M., Mittal, A., & Vishwakarma, D. K. , 2021) mengusulkan metode klasifikasi yang berfokus pada fitur audio dari cuplikan dapat dilakukan sangat baik dengan total data latih 78 fitur yang telah diekstraksi. Dalam hal ini, audio diekstraksi menjadi potongan-potongan kecil agar mendapatkan karakteristik suara tersebut diantaranya untuk memperoleh gelombang sebagai bentuk koordinat. Selanjutnya peneliti yang dilakukan (Huang, Y. F., & Wang, S. H. , 2012) menggunakan algoritma optimasi *metaheuristic* atau *Self-Adaptive Harmony Search* (SAHS) untuk memilih fitur

yang dapat membedakan setiap genre sehingga mendapatkan subset fitur yang lebih baik. Kedua peneliti ini menggunakan dataset film barat.

Berdasarkan uraian diatas dan penelitian yang telah dilakukan maka penelitian akan menggunakan Metode *Support Vector Machine* (SVM) untuk melakukan klasifikasi genre film berdasarkan cuplikan dan melakukan pengembangan terhadap algoritma-algoritma yang sudah diterapkan sebelumnya oleh beberapa penelitian. Serta memberikan usul untuk melakukan pengklasifikasian pada cuplikan film Indonesia ke dalam lima genre film berdasarkan cuplikan film dengan menggunakan Metode *Support Vector Machine* (SVM) dengan harapan bahwa akan mendapatkan hasil yang maksimal memberikan rekomendasi algoritma terbaik.

Penelitian ini berfokus pada klasifikasi dan identifikasi Video genre dari cuplikan (trailer). Penelitian ini dilakukan berdasarkan *low-level and high level features extracted* dari saluran audio yang diekstraksi dari cuplikan film, sehingga metode yang diajukan menggabungkan pendekatan audio dengan metode berbasis gambar sederhana dalam sistem audio-video. Pendekatan yang peneliti usulkan untuk klasifikasi genre (*trailer*) adalah dengan menggunakan fitur audio yang berasal dari video trailer yang berdurasi kurang lebih 3 (tiga) menit.

1.2 Rumusan dan Batasan Masalah Penelitian

Berdasarkan pada latar belakang masalah penelitian yang telah diuraikan dan sesuai dengan usulan topik penelitian, maka dirumuskan tiga masalah pokok yang harus diteliti diantaranya sebagai berikut :

1. Bagaimana mengekstraksi dan merepresentasikan fitur-fitur suara yang ada pada video atau cuplikan (*trailer*) untuk digunakan dalam proses pengelompokan atau identifikasi dari kelima genre.
2. Bagaimana menentukan metode klasifikasi yang tepat untuk mengidentifikasi genre film berdasarkan dari cuplikan (*trailer*) agar mendapatkan nilai akurasi yang maksimal dalam mengkategorikan genre film.
3. *Dataset* yang dipakai merupakan video yang telah di filter suaranya yang berasal dari sumber Kaggle.com

Cuplikan (*trailer*) penelitian dalam kualifikasi ini dibatasi hanya lima genre yaitu thriller, horor, aksi, drama/romantis dan komedi. Dalam hal ini hanya film Indonesia saja yang akan diklasifikan.

1.3 Tujuan Penelitian

Berdasarkan masalah penelitian yang telah diuraikan tersebut, maka tujuan yang ingin dicapai dalam penelitian adalah sebagai berikut :

1. Menghasilkan MFCC (*Mel-Frequency Cepstrum Coefficients*) yang dapat mengidentifikasi ciri dari masing-masing cuplikan (*trailer*)
2. Menghasilkan algoritma pengklasifikasian genre berdasarkan dari cuplikan (*trailer*) film dengan menggunakan metode *Support Vector Machine* (SVM).
3. Mendapatkan kumpulan dataset yang mewakili semua fitur agar kelengkapan data trainer terpenuhi.

1.4 Kontribusi Penelitian

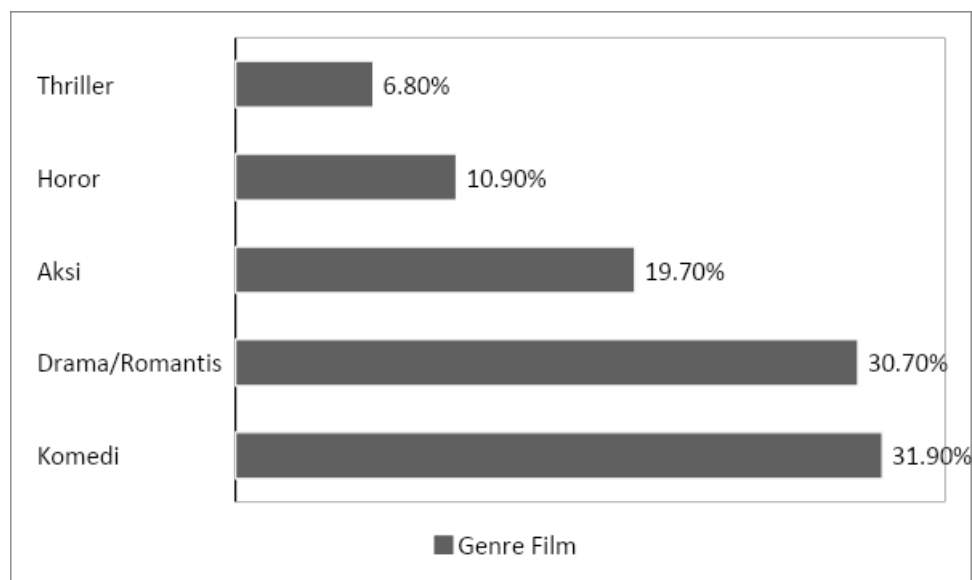
Manfaat dan kontribusi dari penelitian ini adalah :

1. Dari segi keilmuan, penelitian ini akan memberikan kontribusi berupa pengembangan metode dan teknologi pengklasifikasian suara yang mengandung informasi teks berbahasa indonesia.
2. Dari sisi metode, penelitian ini diusahakan untuk menemukan cara baru/penambahan/modifikasi dari metode yang ada untuk proses pemfilteran suara yang inputannya berupa video yang kemudian akan diterjemahkan menjadi teks dalam satu proses. Dimana teks tersebut harus melewati proses ekstraksi fitur suara agar memperoleh ciri fitur yang nantinya akan masuk menjadi Data Set.

2. TINJAUAN PUSTAKA

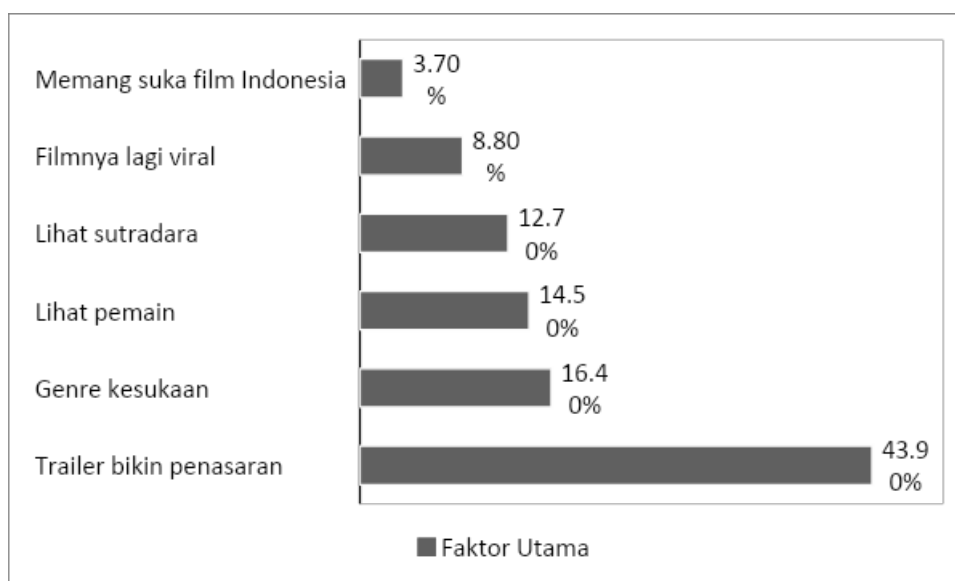
2.1 Tren Film Di Indonesia

Film merupakan karya sinematografi yang dapat berfungsi sebagai alat pendidikan budaya. Dengan demikian film juga efektif untuk menyampaikan nilai-nilai budaya (Trianton, 2013). Pilihan dan perbedaan penonton merupakan pengaruh kuat dalam persaingan media, terutama media audio dan visual. Film Indonesia mulai bangkit kembali pasca gerakan reformasi tahun 1998. Berawal dari menurunnya semangat dalam produksi film Indonesia pada pertengahan tahun 90-an. Berdasarkan data IDN Times pada tahun 2019 bahwa pemilihan film di Indonesia dapat dilihat berdasarkan dari genre film tersebut. Gambar 2.1 merupakan persentase genre terfavorit pada film di Indonesia yang berasal dari 411 responden (Film Indonesia, 2010). Oleh karena itu, peneliti akan mengambil lima genre dari empat belas genre yang ada.



Gambar 2.1 Infografis yang diminati berdasarkan genre (Sumber : IDN Times, 2019)

Sumber informasi selanjutnya mengenai perkembangan faktor utama peminat film di Indonesia tahun 2019 dapat dilihat pada gambar 2.2



Gambar 2.2 Infografis Minat Film Pada Tahun 2019 (Sumber : IDN Times, 2019)

Sumber gambar 2.1 dan gambar 2.2 berasal dari perusahaan media multi-platform untuk generasi Milenial dan Gen Z di Indonesia yaitu IDN Times sejak 8 Juni 2014 (IDN Times, 2019). Berdasarkan hasil riset IDN Media (2019), generasi milenial banyak mengakses dan menerima informasi dari media digital. Menurut sumber, sebesar 43,9% menunjukkan bahwa Cuplikan (*trailer*) dapat meningkatkan minat penonton karena di dalamnya memuat nama pemain yang membintangi film tersebut. Banyaknya cuplikan adegan yang ditampilkan dalam *trailer* memiliki durasi sekitar 2-3 menit. Berawal dari fenomena diatas, peneliti bermaksud untuk melakukan klasifikasi genre film berdasarkan dari cuplikan (*trailer*).

Cuplikan (*trailer*) berisikan adegan dari sebuah film secara lebih lengkap dan urut yang disajikan dalam bentuk video yang dapat memberikan sedikit gambaran mengenai alur cerita dan konflik yang ada pada film tersebut. Dalam video cuplikan ini akan menampilkan bagian awal, tengah, hingga akhir film secara garis besar.

2. 2 Pengenalan Video dan Audio (Suara)

Menurut Sukiman dan Daryanto dalam teorinya menyampaikan bahwa video merupakan sebuah media yang bisa memperlihatkan gambar, disertai suara, yang tampil pada saat yang bersamaan yang didalamnya terdapat suara (*audio*), teks yang dapat direkam melalui media digital agar menghasilkan sebuah putaran salah satunya adalah film.

Setiap video seharusnya memiliki kelas yang unik dan menjadi satu cerita, yaitu struktur naratifnya relatif sederhana dan homogen (Rouvier, M, 2007). Video menyajikan gambar bergerak dengan suara yang menyertainya.

Pengenalan suara (*voice recognition*) dibagi menjadi dua jenis, yaitu *speech recognition* dan *speaker recognition*. *Speech recognition* adalah proses yang dilakukan komputer untuk mengenali kata yang diucapkan oleh seseorang tanpa mempedulikan identitas orang terkait. *Speaker recognition* merupakan pengenalan identitas yang diklaim oleh seseorang dari suaranya (ciri khusus dapat berupa intonasi suara, tingkat kedalaman suara dan sebagainya).

Untuk memungkinkan terjadinya pengenalan suara yang baik, maka pertama-tama sistem melakukan penyimpanan suatu database untuk mengetahui setiap arti dari kata yang diberikan padanya. Terdapat beberapa faktor yang perlu diperhatikan, yaitu kondisi lingkungan, jumlah kosa kata, kemiripan kosakata, dan lafal pengucapan kata (Shafirra, N. A, dkk, 2020). Audio dibagi menjadi potongan-potongan kecil yang ada pada video. Secara umum, sinyal audio dapat diukur dengan dua cara berbeda yaitu domain waktu dan domain frekuensi (Huang, Y. F., & Wang, S. H. , 2012)

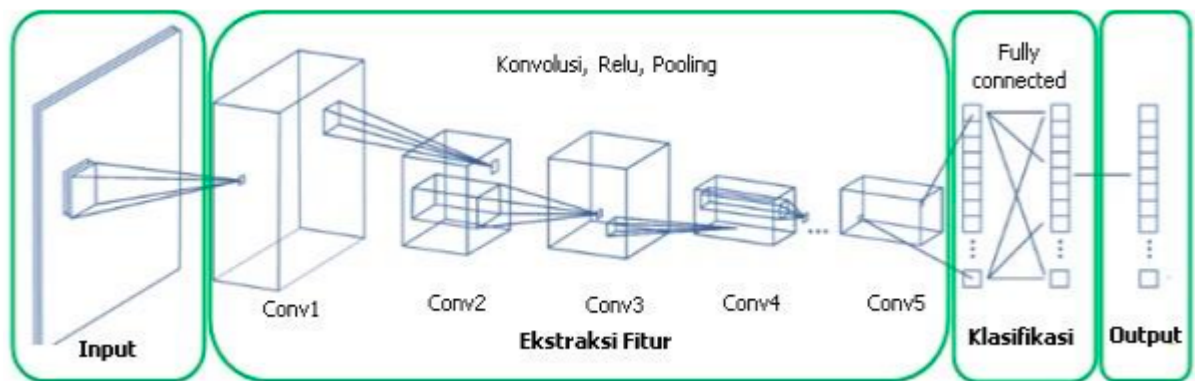
2. 3 Klasifikasi

Klasifikasi merupakan cara pengelompokkan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual maupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya *Convolutional Neural Network* dan *Support Vector Machine* (Aji Prasetya Wibawa, 2018). Sebuah pengklasifikasi dibuat dari sekumpulan data latih dengan kelas yang telah ditentukan (M. A. Shadiq, 2009) sehingga dapat membantu memahami keragaman dengan lebih baik karena klasifikasi merupakan suatu proses yang berkaitan dengan kategorisasi. Sehingga, klasifikasi merupakan pengelompokan fitur ke dalam kelas yang sesuai dan proses identifikasi kelompok dari suatu objek berdasarkan kesamaan fitur tertentu, dimana setiap kelompok telah terbentuk melalui suatu proses tertentu. Keunikan setiap kelompok direpresentasikan dengan suatu vektor fitur (R. Y. Dillak, 2022).

2.3.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah algoritma *deep learning* yang dapat menggunakan gambar sebagai input, menetapkan kepentingan (bobot dan bias yang dapat dipelajari) untuk berbagai aspek dan objek dalam gambar agar dapat membedakan satu dengan yang lain (Trianton, T, 2013).

Struktur CNN terdiri dari input, proses ekstraksi fitur, proses klasifikasi dan output. Proses ekstraksi dalam CNN terdiri dari beberapa lapisan tersembunyi atau hidden layer, yaitu lapisan konvolusi, fungsi aktivasi (ReLU), dan pooling. CNN bekerja secara hierarki, sehingga output pada lapisan konvolusi pertama digunakan sebagai input pada lapisan konvolusi selanjutnya. Pada proses klasifikasi terdiri dari fully-connected dan fungsi aktivasi (*softmax*) yang outputnya berupa hasil klasifikasi (Katole et al., 2015).

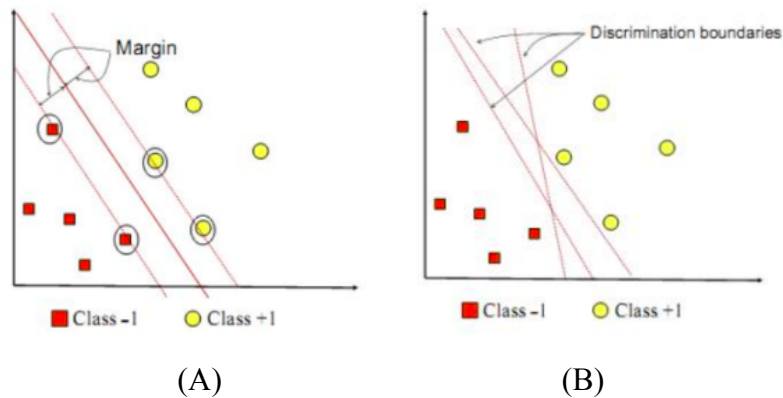


Gambar 2.1 Arsitektur *Convolutional Neural Network* (Sumber : Sumber :Krizhevsky et al. , 2012)

Fungsi aktivasi *softmax* digunakan untuk mendapatkan hasil klasifikasi. Fungsi aktivasi menghasilkan nilai yang diinterpretasi sebagai probabilitas yang belum dinormalisasi untuk tiap kelas. Nilai kelas dihitung dengan menggunakan fungsi *softmax* (Vedaldi & Lenc, 2015).

2.3.2 Support Vector Machine (SVM)

SVM adalah metode *learning machine* yang bekerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah class pada input space. *Hyperplane* pemisah terbaik antara kedua class dapat ditemukan dengan mengukur margin *hyperplane* tersebut dan mencari titik maksimalnya (C. Y. Feriyawan, 2012). Usaha untuk mencari lokasi *hyperplane* merupakan inti dari proses pembelajaran pada SVM. Klasifikasi SVM dilakukan dengan cara menemukan *hyperplane* terbaik sehingga diperoleh ukuran margin yang maksimal. Margin adalah jarak antara *hyperplane* tersebut dengan titik terdekat dari masing-masing kelas. Titik yang paling dekat ini disebut dengan *support vector*.



Gambar 2.2 (A) Sebelum terbentuk (B) Sesudah terbentuk diantara kedua Class
(Sumber : Aji Prasetya Wibawa, 2018)

Berdasarkan ilustrasi gambar 2.2 , mengilustrasikan dua kelas yang dapat dipisahkan secara linier menggunakan sepasang bidang batas yang sejajar. Kemudian Bidang pemisah terbaik adalah bidang pemisah yang memiliki margin (jarak antara dua bidang pembatas) maksimal atau berada di tengah-tengah kedua kelas yang berbeda.

2.4 Perkembangan Penelitian Terkait

2.4.1 Implementasi Metode *Convolutional Neural Network* (CNN) Pada Klasifikasi Video

Penelitian yang dilakukan oleh Gabriel S. Simoes dan kawan - kawan (Simões, 2016) adalah melakukan klasifikasi genre film ke dalam empat yaitu aksi, komedi, horor, atau drama dengan menerapkan CNN Mo-Tion (*Convolutional Neural Networks for Movie Trailer Classification*) untuk menganalisis video berdasarkan *Learning from low-level features* seperti , seperti rata-rata panjang bidikan, varian warna, tombol pencahayaan, dan kehadiran gerakan (konten gerak) dan *Learning from high-level features* yaitu untuk klasifikasi genre film menggunakan deskriptor gambar yang telah dikenal untuk menghitung fitur tingkat tinggi yang ada pada setiap bingkai utama sehingga memperoleh similaritas dari informasi semantik seperti kealamian, keterbukaan, kekasaran, ekspansi, dan kekasaran yang mewakili struktur spasial yang dominan dari sebuah adegan. Hal ini bertujuan untuk mengenali tempat dan pemandangan. Untuk menghasilkan CENTRIST *features*.

Hasil analisis eksperimental pada penelitian tersebut bahwa CNN-MoTion mengungguli semua pendekatan berbasis gambar/video. Berdasarkan ukuran akurasi, variasi CNN-MoTion terbaik adalah 73,45%. Untuk detail kinerja lebih lanjut, metode CNN memiliki kekhususan ($TN/(TN+FP)$) sebesar 97,41% untuk memprediksi film horor. Untuk

komedi, adalah 89,13%. Namun, film aksi yang diprediksi sebagai drama, yang terjadi 14 kali dan film horor dan drama diprediksi sebagai trailer aksi.

2.4.2 Implementasi Metode *Support Vector Machine* (SVM) Pada Klasifikasi Video

Penelitian yang dilakukan oleh Vakkalanka Suresh dan kawan - kawan (Suresh, 2004) adalah melakukan klasifikasi lima video diantaranya kartun, komersial, cricket, sepak bola dan tenis dengan menggunakan pendekatan yang didasarkan pada spasial dan temporal *descriptor* yang berasal dari video pendek (20 detik) yang menggunakan Metode *Support Vector Machine* (SVM) telah berhasil membedakan antara kelas pada fitur yang hanya dapat diekstraksi dari konten visual pada video saja yaitu fitur audio, visual dan teks. *Pre-Processing* yang dilakukan untuk ekstraksi fitur diantaranya warna, bentuk, dan gerak. Eksperimen dilakukan pada lebih dari 2 jam data video (≈ 400 klip video dengan masing-masing 20 detik ditangkap pada 25 frame per detik). Hasil yang diperoleh keakurasiannya sebesar 92,5%. Angka ini menunjukkan bahwa kumpulan fitur yang dipertimbangkan dapat memberikan informasi yang berguna untuk pemahaman konten semantik dan memberikan informasi pembeda di antara kelas.

2.5 Perbandingan Penelitian

Penelitian-penelitian terkait pencarian dokumen dengan Penelitian pengenalan suara, teks dan klasifikasi serta film Indonesia yang mengandung informasi cuplikan disajikan pada tabel 2.1.

Tabel 2.1 Ringkasan Penelitian Klasifikasi

No	Penelitian, tahun	Ringkasan hasil	Kaitan dengan penelitian	Celah yang ditemukan (peluang)
1	Metode Metode Klasifikasi [Aji Prasetya Wibawa, dkk , 2018]	Jaringan Saraf Tiruan (Neural Network), Naïve Bayes, Support Vector Machine, Decision Tree, dan Fuzzy	Sebagai pustaka acuan bahwa untuk mencari tahu metode apa saja yang digunakan untuk mengklasifikasikan data sesuai dengan ciri-ciri atau karakteristik dari data yang dimiliki.	Proses untuk menentukan pencarian metode apa yang akan digunakan serta mengidentifikasikan masalah mengenai faktor apa saja untuk mendapatkan ciri atau kategori pada data yang akan digunakan (suara).

No	Penelitian, tahun	Ringkasan hasil	Kaitan dengan penelitian	Celah yang ditemukan (peluang)
2	Klasifikasi Film Berdasarkan Sinopsis dengan Menggunakan Improved K-Nearest Neighbor (K-NN) [Nurul Muslimah, dkk, 2019]	Dengan menggunakan metode Improved-KNN (K-Nearest Neighbor) klasifikasi berdasarkan dokumen yang berupa sinopsis mendapatkan hasil terbaik yakni precision =1; recall=0,88; f-measure=0,936170213 yang menggunakan 250 dokumen (data latih) dan 50 dokumen data uji. Dimana tingkat akurasi 88%.	Sebagai landasan awal bahwa klasifikasi film dapat ditentukan dengan menggunakan sinopsi yang mana akan dipersempit lagi dengan menggunakan cuplikan.	Belum mendapatkan acuan mengenai bagaimana mengkonversikan suara ke dalam bentuk teks terutama video pada cuplikan yang memiliki efek suara.
3	Klasifikasi Sentimen Ulasan Film Indonesia Dengan Konversi Speech-To-Text (Stt) Menggunakan Metode Convolutional Neural Network (Cnn) [Nadhifa Ayu Shafirradan Irhamah, 2020]	Metode yang digunakan adalah Convolutional Neural Network (CNN) dimana pada penelitian ini menggunakan konvolusi 6 layer dengan ukuran filter sebesar 1, 2 dan 3. Kelebihan	Sebagai landasan awal bahwa video dapat dikonversikan ke dalam bentuk teks dengan bantuan STT penentu metode yang baik dalam hal mengklasifikasikan <i>trailer</i> film yang mana akan diklasifikasikan berdasarkan <i>gender</i> .	Menemukan penanganan untuk melakukan konversi suara ke dalam bentuk teks namun belum teruji untuk suara yang memiliki noise (efek) yang ada pada cuplikan film.
4	Speech Word Recognition to Text Converter Using Support Vector Machine [elsen ronando, sugiono, 2019]	Metode <i>linear predictive coding</i> digunakan untuk mendapatkan ekstraksi ciri dari set data (telah menghasilkan 2 bagian set) yaitu data set latih dan data set uji. Dimana untuk melatih ke-2 data menggunakan Support Vector Machine untuk mengenali ucapan	Sebagai landasan dan mencari metode lain yang tepat untuk mendapatkan ekstraksi ciri dari data set yang akan digunakan.	Dapat mengklasifikasikan data secara linier atau non-linier.

No	Penelitian, tahun	Ringkasan hasil	Kaitan dengan penelitian	Celah yang ditemukan (peluang)
		kata ke dalam bentuk teks dengan akurasi 71,875%		
5	Metode Mel <i>Frequency Cepstral Coefficients</i> (MFCC) Pada klasifikasi Hidden Markov Model (HMM) Untuk Kata Arabic pada Penutur Indonesia [Totok Chamidy, 2016]	Metode Mel <i>Frequency Cepstral Coefficients</i> digunakan untuk mendapatkan ekstraksi ciri dari suara menjadi kata. Dimana untuk mengklasifikasi dan mendapatkan fitur yang telah diekstrak menggunakan Metode Hidden Markov Model dengan akurasi sebesar 83,1% untuk frekuensi sampling data uji sebesar 8000 Hz, 82,3% untuk frekuensi sampling data uji sebesar 22050 Hz, 82,2% untuk frekuensi sampling data uji sebesar 44100Hz. Dengan 128 data suara (data latih) dan 32 data suara (data uji)	Sebagai landasan awal bahwa untuk meningkatkan akurasi harus menambahkan data latihnya.	Diharapkan bisa mengenali ribuan kata dan mengetahui akurasi klasifikasi dari HMM untuk mengenali ucapan dalam satu frekuensi sampling.
6	Audio-based video genre identification [Mickael Rouvier, Stanislas Oger, Georges Linare's, Driss Matrouf, Bernard Merialdo, Yingbo Li, 2007]	Dengan Sistem berbasis audio-saja terbaik melakukan 9% dari Tingkat Kesalahan Klasifikasi (CER) dan untuk mengevaluasi kelengkapan fitur audio yang diusulkan adalah dengan fitur video yang digunakan secara klasik untuk Identifikasi Genre Video (VGI). Hasil menunjukkan komplementaritas	Sebagai informasi bahwa audio tampaknya secara signifikan kurang akurat daripada video untuk dilakukan identifikasi genre.	Dapat memperkuat bahwa klasifikasi yang akurat adalah dengan menggunakan video namun harus menggunakan metode yang relevan agar hasilnya optimal.

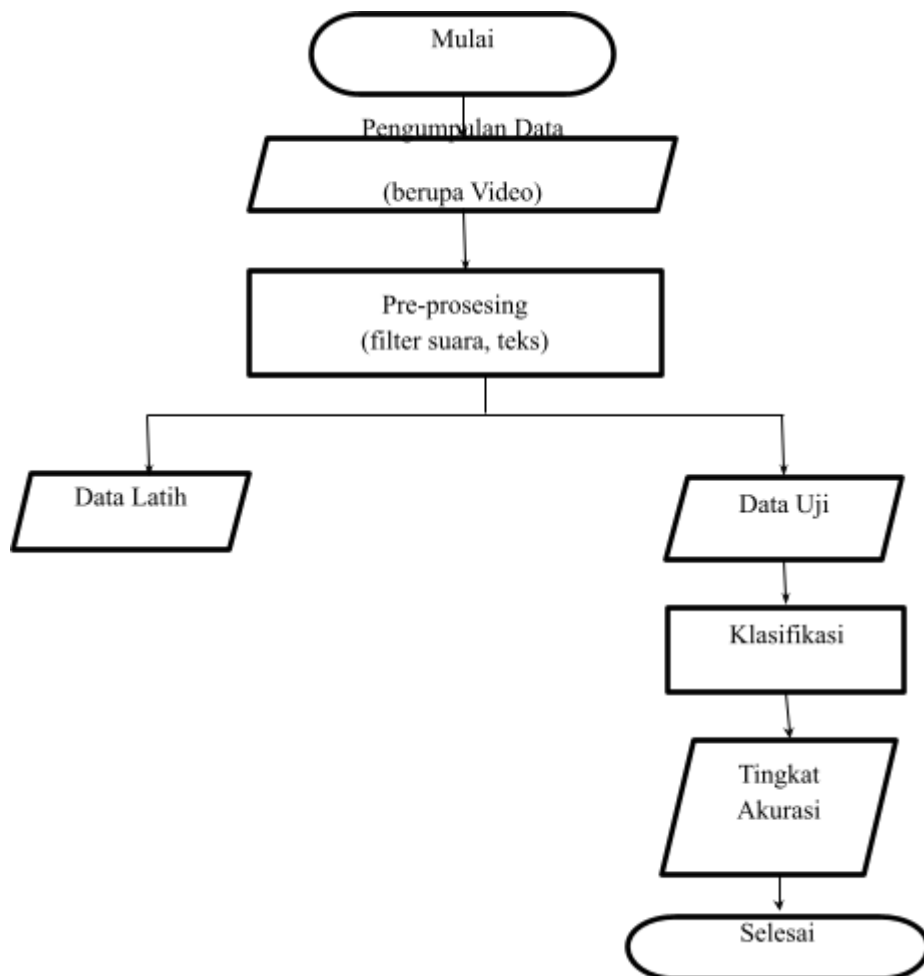
No	Penelitian, tahun	Ringkasan hasil	Kaitan dengan penelitian	Celah yang ditemukan (peluang)
		modalitas untuk pengenalan genre, sistem audio-video akhir mencapai 6% CER.		
7	Automatic Video Genre Categorization Using Hierarchical [Xun Yuan, Wei Lai, Tao Mei, Xian-Sheng Hua, Xiu-Qing Wu, Shipeng Li, 2006]	Telah berhasil melakukan kategorisasi genre video otomatis dengan menggunakan fitur tingkat rendah spasial-temporal dengan menggunakan skema SVM namun berdasarkan dari audio, gerakan kamera, pencahayaan (lampu kilat).	Sebagai aturan mengenai pelabelan genre video dengan menggunakan vektor fitur dimensi yang belum ditentukan banyaknya fitur.	Dapat menemukan fitur visual untuk genre 3 (tiga) dari 5 (lima) movie yang akan diteliti yaitu Aksi, Komedi, horor.
8	Content-Based Video Classification Using Support Vector Machines [Vakkalanka Suresh, C. Krishna Mohan, R. Kumara Swamy, and B. Yegnanarayana, 2004]	Pendekatan yang dilakukan untuk masalah klasifikasi video kedalam genre didasarkan pada <i>spatial</i> dan temporal <i>descriptors</i> yang berasal dari cuplikan video pendek (20 detik) dan mencapai akurasi 92,5%	Sebagai penambahan ilmu bahwa fitur visual saja mungkin tidak cukup sehingga modalitas audio dan teks akan digabungkan dengan bukti visual agar cocok dengan genre.	Dapat mengetahui ekstraksi fitur visual spasial dan temporal yang melekat dalam konten video / cuplikan (fitur warna, bentuk, dan gerakan).
9	A Unified Audio Analysis Framework For Movie Genre Classification Using Movie Trailers [Aditya Sharma, Mayank Jindal, Dinesh Kumar Viswakarma, 2021)	Dengan pendekatan <i>unsupervised learning</i> memudahkan proses klasifikasi genre dengan mempertimbangkan audio dari cuplikan film. Dan menggunakan LMTD-9 untuk tujuan validasi.	Sebagai kerangka kerja dan arsitektur yang akan digunakan dalam penelitian ini	Dapat memprediksi genre film dengan menggunakan pendekatan pembelajaran <i>unsupervised learning</i> dan menggunakan usulan jurnal ini sebagai kerangka kerja untuk klasifikasi genre.
10	Movie Genre Classification Using SVM with Audio and Video Features	Dengan menggunakan algoritma optimasi meta-heuristik yang disebut <i>Self-Adaptive</i>	Sebagai algoritma optimasi yang akan dikembangkan dalam penelitian ini	Menghasilkan subset fitur untuk semua genre film dan mendapatkan fitur yang optimal dari set fitur asli.

No	Penelitian, tahun	Ringkasan hasil	Kaitan dengan penelitian	Celah yang ditemukan (peluang)
	(Huang, Y. F., & Wang, S. H., 2012)	<i>Harmony Search</i> (yaitu, SAHS) untuk memilih fitur lokal untuk genre film yang sesuai maka telah didapat hasil pencapaian sebesar 91,9% untuk setiap pasangan genre untuk mendapatkan hasil klasifikasi yang lebih baik.		
11	Pengklasifikasian Genre Musik Indonesia Menggunakan Convolutional Neural Network [Wairata, C. R., Swedia, E. R., & Cahyanti, M, 2021]	Untuk penggunaan CNN dalam data audio ini, yang dilakukan adalah mengubah suara menjadi sebuah gambar dengan koordinat x sebagai Amplitudo dari audio itu sendiri, sedangkan koordinat y sebagai durasi waktu dari audio tersebut. Percobaan ke-10 dengan menggunakan 35 epoch telah mendapatkan akurasi validasi sebesar 100% dan akurasi tes sebesar 81,33%.	Sebagai penambahan ilmu bahwa dataset musik dapat menghasilkan gelombang suara dari lagu.	Metode MFCC ini dapat digunakan dalam penelitian.

3. METODOLOGI PENELITIAN

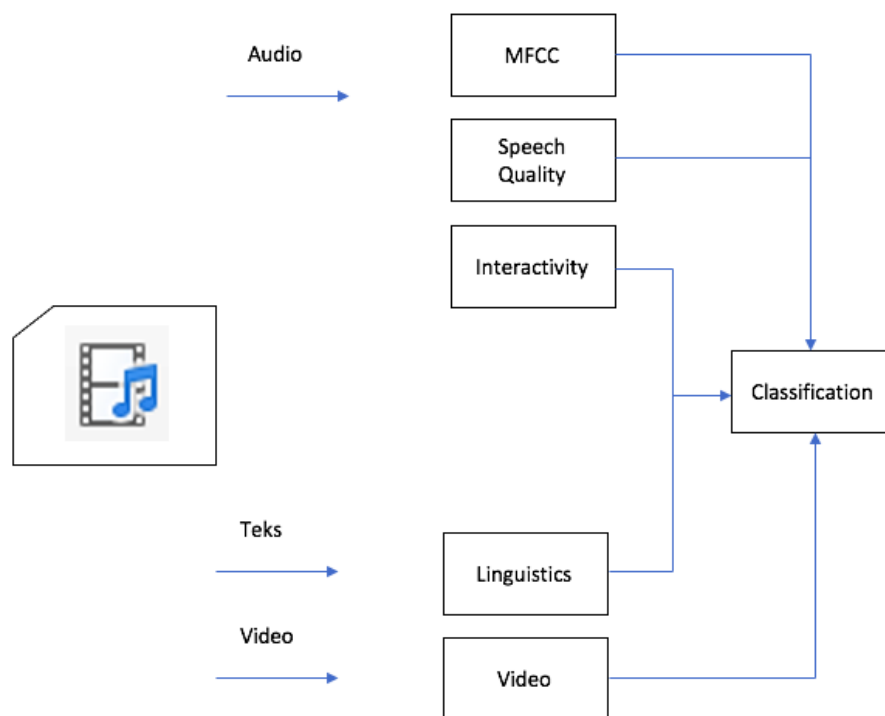
3.1 Tahapan Penelitian

Penelitian ini berusaha untuk mendapatkan klasifikasi genre film di Indonesia berdasarkan dari penayangan film atau dikenal dengan istilah cuplikan (*thriller*) yang dapat membantu para peminat film agar dapat membedakan genre Komedi, Drama/Romantis, Aksi, Horor dan Thriller. Rencana penelitian mencoba untuk menggabungkan teknik dan mengembangkan algoritma yang dapat menghasilkan klasifikasi film ke dalam lima genre Komedi, Drama/Romantis, Aksi, Horor dan Thriller sehingga didapat metode yang menghasilkan nilai akurasi yang maksimal. Berdasarkan peneliti sebelumnya, untuk mendapatkan data latih dan uji maka tahapan terpentingnya adalah pada tahapan preprocessing. Pada tahapan awal ini, maka data yang akan dilatih dan diuji dapat dilakukan. Berikut tahapan peneliti dalam mengklasifikasikan lima genre yang di dapat dari inputan video yang akan diuraikan pada gambar 3.1 dibawah ini.



Gambar 3.1 Tahapan Penelitian Yang Akan Dilakukan

Peneliti akan menggunakan pendekatan yang sama pada penelitian (Mickael Rouvier, dkk, 2007) untuk identifikasi suara (audio) yang ada pada cuplikan (*trailer*) dengan cara mengambil dari *background* musik tersebut namun berbeda dataset, yaitu Film Indonesia saja. Diharapkan dapat menghasilkan fitur audio tingkat tinggi dengan tujuan menangkap informasi relatif terhadap konteks audio dan struktur temporal video. Dengan menggunakan metode *Support Vector Machine* (SVM). Metode Klasifikasi yang akan dilakukan pada penelitian ini adalah pada gambar berikut :



Gambar 3.2 Metode Klasifikasi (Sumber : Rouvier, M, 2015)

Metode klasifikasi yang tertera pada gambar 3.2 adalah sumber yang berupa video atau cuplikan film, penelitian dilakukan dengan memisahkan sumber dari video yaitu, audio, text dan video untuk mengklasifikasi genre dari film tersebut. Audio ini akan diekstraksi dari trailer video. Detail dari metode sebagai berikut :

1. MFCC Cepstral, merupakan pola pemodelan statistik yang berasal dari audio, untuk kategori berdasarkan audio atau frekuensi sinyal audio per waktu untuk dapat identifikasi ciri dari suatu cuplikan film
2. *Speech quality*, pada metode ini berdasarkan pada pembicaraan atau pengenalan suara untuk memperkirakan genre film berdasar kata kata yang diucapkan, dalam metode ini juga kualitas suara mempengaruhi dari identifikasi

3. *Interactivity*, pada bagian ini sistem identifikasi percakapan antara satu orang dengan lainnya dalam cuplikan film, inti dari percakapan dapat mengetahui genre dari video tersebut
4. *Linguistic*, pada bagian ini identifikasi dilakukan dengan melihat teks pada film (jika ada), pada bagian linguistic erat kaitanya dengan interactivity
5. Video, pada bagian video, genre dapat diidentifikasi dari background film, ataupun ekspresi muka dan perilaku.

Kelima metode tersebut dapat dijadikan kombinasi untuk klasifikasi genre dari film. Dalam hal ini, peneliti akan melakukan 2 (dua) fase, pertama mengambil suara (*audio*) dari *background* musik cuplikan dan fase kedua menangkap *audio speech quality* berupa percakapan dan akan digabung dengan *linguistik* dan interaktif isi dari percakapan dalam trailer tersebut.

III.3. RENCANA KERJA

Tabel 3.1. Rencana Pelaksanaan Penelitian

Kegiatan	Bulan Ke-																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Studi Pustaka																	
Pemilihan Topik																	
Penyusunan Proposal																	
Pengajuan Proposal																	
Pengerjaan Penelitian																	
Publikasi Ilmiah / Seminar																	
Penyusunan Desertasi																	
Sidang Desertasi																	

DAFTAR PUSTAKA

- B. Syahid, "Pengertian Teks Ulasan, Contoh, Ciri, Tujuan, Struktur, dan Kaidahnya," guru pendidikan, Jan. 13, 2020. <https://www.gurupendidikan.co.id/teks-ulasan/> (accessed Des. 12, 2020).
- Chamidy, T. (2016). Metode Mel *Frequency Cepstral Coefficients* (MFCC) Pada klasifikasi Hidden Markov Model (HMM) Untuk Kata Arabic pada Penutur Indonesia. *Matics*, 8(1), 36-39.
- C. Y. Feriyawan and P. Danoedoro, "Kajian Kemampuan Jaringan Syaraf Tiruan Algoritma Backpropagation untuk Klasifikasi Penggunaan Lahan Menggunakan Citra ALOS AVNIR-2," J. Bumi Indones., vol. 1, no. 3, pp. 101–110, 2012.
- Film Indonesia, (2010). <http://filmindonesia.or.id/movie/trailer#.X9n5SLOyTIU> (accessed Des. 16, 2020).
- Huang, Y. F., & Wang, S. H. (2012, December). Movie genre classification using svm with audio and video features. In *International Conference on Active Media Technology* (pp. 1-10). Springer, Berlin, Heidelberg.
- Katole, A. L., Yellapragada, K. P., Bedi, A. K., Kalra, S. S., & Siva Chaitanya, M. (2015). Hierarchical Deep Learning Architecture for 10K Objects Classification. *Computer Science & Information Technology (CS & IT)*, (September), 77–93.
- M. A. Shadiq, "Keoptimalan Naive Bayes dalam Kualifikasi," Universitas Pendidikan Indonesia, 2009.
- Muslimah, N., & Indriati, R. C. W. (2019). Klasifikasi film berdasarkan sinopsis dengan menggunakan improved K-Nearest Neighbor (K-NN). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X*.

R. Y. Dillak, D. M. Pangestuty, and M. G. Bintiri, "Klasifikasi Jenis Musik Berdasarkan File Audio Menggunakan Jaringan Syaraf Tiruan Learning Vector Quantization," in Seminar Nasional Informatika, 2012, pp. 122–125.

Ronando, E. (2019). Sistem Konversi Ucapan Kata Ke Teks Menggunakan Support Vector Machine: Speech Word Recognition To Text Converter Using Support Vector Machine. *Jurnal Teknologi dan Terapan Bisnis*, 2(2), 1-8.

Rouvier, M., Oger, S., Linares, G., Matrouf, D., Merialdo, B., & Li, Y. (2015). Audio-based video genre identification. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(6), 1031-1041.

Shafirra, N. A., & Irhamah, I. (2020). Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Speech-to-Text (STT) Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Sains dan Seni ITS*, 9(1), D95-D101.

Sharma, A., Jindal, M., Mittal, A., & Vishwakarma, D. K. (2021, March). A Unified Audio Analysis Framework For Movie Genre Classification Using Movie Trailers. In *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)* (pp. 510-515). IEEE.

Simões, G. S., Wehrmann, J., Barros, R. C., & Ruiz, D. D. (2016, July). Movie genre classification with convolutional neural networks. In *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 259-266). IEEE.

Suresh, V., Mohan, C. K., Swamy, R. K., & Yegnanarayana, B. (2004, November). Content-based video classification using support vector machines. In *International conference on neural information processing* (pp. 726-731). Springer, Berlin, Heidelberg.

Trianton, T. (2013). Film Sebagai Media Pelajaran. Yogyakarta: Graha Ilmu

Wairata, C. R., Swedia, E. R., & Cahyanti, M. (2021). PENGKLASIFIKASIAN GENRE MUSIK INDONESIA MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK. *Sebatik*, 25(1), 255-261.

Wibawa, A. P., Purnama, M. G. A., Akbar, M. F., & Dwiyanto, F. A. (2018, April). Metode-metode Klasifikasi. In *Prosiding SAKTI (Seminar Ilmu Komputer dan Teknologi Informasi)* (Vol. 3, No. 1, pp. 134-138).

Widya, D. A., (2017). Sinematografi. Universitas Dian Nuswantoro. Fakultas Ilmu Komputer