

## Klasifikasi Video Olahraga Berdasarkan Citra Berbasis Konten Menggunakan Segmentasi *Superpixel*

Yohannes<sup>1\*</sup>, Nur Rachmat<sup>2</sup>, Bobby Jaya Saputra<sup>3</sup>

<sup>1,2,3</sup>Program Studi Informatika, Fakultas Ilmu Komputer dan Rekayasa, Universitas Multi Data Palembang  
Jl. Rajawali No. 14, Palembang, Indonesia 30113

\*email: yohannesmasterous@mdp.ac.id

(Naskah masuk: 9 Februari 2021; diterima untuk diterbitkan: 19 Maret 2021)

**ABSTRAK** – Jenis olahraga memiliki jumlah yang banyak dan semakin bertambah dari waktu ke waktu. Banyak orang yang suka menyaksikan perlombaan atau turnamen olahraga di lapangan secara langsung. Berkat kemajuan teknologi, kini turnamen olahraga bisa disaksikan secara digital melalui video tanpa perlu menontonnya secara langsung di lapangan. Klasifikasi video perlu dilakukan untuk membedakan banyaknya video olahraga saat ini ke dalam kategori olahraga yang sesuai dengan kontennya. Penelitian ini mengklasifikasikan video olahraga berjenis Baseball, Basketball, Boxing, Gymnastic, Hockey, Swimming, Tennis, Volleyball, dan Wrestling berdasarkan citra berbasis konten. Klasifikasi dilakukan dengan konten citra sebagai fitur, kemudian dilakukan proses segmentasi *superpixel* dan *superpixel colorization* pada citra untuk membentuk model. Model dibentuk menggunakan metode CNN dengan arsitektur ResNet-50. Berbagai jumlah segmentasi *superpixel* digunakan sebagai perbandingan dalam pembentukan model untuk mendapatkan hasil yang terbaik yaitu 25.000 sehingga digunakan untuk pengujian video olahraga. Hasil pengujiannya yaitu segmentasi *superpixel* bisa digunakan untuk mengklasifikasikan video olahraga dengan baik dengan tingkat rata-rata accuracy 0,91; precision 0,64; dan recall 0,61, walaupun terdapat beberapa olahraga yang memiliki hasil yang kurang baik seperti basketball dan wrestling.

**Kata Kunci** – Klasifikasi Video Olahraga Berbasis Konten; Simple Linear Iterative Clustering; Segmentasi *Superpixel*; *Superpixel Colorization*; ResNet-50.

## Sport Video Classification Based on Content Image with *Superpixel* Segmentation

**ABSTRACT** – Type of sport has a large number and its number is always increasing as the time goes by. Many people like to watch sports competitions or tournaments on the field live. Nowadays, sports tournaments can now be watched digitally via video without the need to watch them live on the field by technological advances. Video classification needs to be done to differentiate the number of sports videos currently into the sports category according to the content. This study classifies sports videos of the types of Baseball, Basketball, Boxing, Gymnastic, Hockey, Swimming, Tennis, Volleyball, and Wrestling based on image content. The classification is done with the image content as a feature, then the *superpixel* segmentation and *superpixel colorization* processes are carried out on the images to form a model. The model is formed using the CNN method with ResNet-50 architecture. Various numbers of *superpixel* segmentation are used as a comparison in the modelling to get the best results. The number of *superpixel* segmentation with the best result is 25.000 so that it is used for testing sports videos. The test result is the *superpixel* segmentation can be used to classify sports videos well with an average rate of accuracy 0.91, precision 0.64 and recall 0.61, although some sports that had poor results, such as basketball and wrestling.

**Keywords** – Sport Video Classification Based-on Content Image; Simple Linear Iterative Clustering; *Superpixel* Segmentation; *Superpixel Colorization*; ResNet-50.

## 1. PENDAHULUAN

Olahraga bermanfaat untuk meningkatkan daya tahan tubuh terhadap sumber penyakit [1]. Selain itu, olahraga juga menjadi hobi bahkan mampu membentuk komunitas dengan hobi yang sama. Olahraga tidak hanya melibatkan fisik, tetapi juga taktik dan strategi. Jumlah olahraga semakin bertambah dari waktu ke waktu. Olahraga ada yang dimainkan secara individu seperti catur, ada juga yang beregu seperti bola basket.

Saat ini, olahraga tidak hanya dilakukan untuk menjaga kesehatan tetapi juga digunakan sebagai ajang perlombaan atau turnamen. Turnamen ini tentu membuat orang atau klub yang sering memenangkan pertandingan memiliki penggemar. Penggemar olahraga tentu tidak ingin kelewatan menonton pertandingan orang atau klub idola, tetapi apabila klub tersebut bertanding di tempat yang jauh menyebabkan tidak bisa menonton secara langsung di lapangan atau stadion. Tetapi, video olahraga saat ini sudah bisa disaksikan secara *online* melalui televisi atau perangkat yang terhubung dengan internet. Video olahraga pun bisa ditonton tidak hanya di siaran langsung namun juga ada yang menyediakan siaran ulang yang bisa ditonton berkali-kali. Klasifikasi video perlu dilakukan untuk membedakan banyaknya video olahraga saat ini ke dalam kategori olahraga yang sesuai dengan kontennya.

Perkembangan teknologi bisa melakukan pengklasifikasian suatu objek termasuk juga jenis video. Video adalah kumpulan dari *frame* dimana setiap *frame*-nya merupakan citra yang bisa diolah untuk berbagai tujuan seperti klasifikasi dan identifikasi suatu objek [2]. *Convolutional Neural Network* (CNN) merupakan salah satu metode dari *Deep Learning* yang cocok untuk melakukan pengenalan dan klasifikasi citra [3]. Metode CNN juga telah digunakan pada banyak penelitian.

Penelitian mengenai klasifikasi video olahraga telah dilakukan sebelumnya oleh beberapa peneliti [4], [5], [6], [7]. Penelitian yang dilakukan oleh [3] mengenai klasifikasi video olahraga menggunakan metode *Convolutional Neural Network* (CNN) menghasilkan akurasi sebesar 92,68% menggunakan arsitektur *AlexNet*, akurasi sebesar 91,67% menggunakan *GoogleNet*, dan akurasi sebesar 88,75% untuk CNN biasa. Selain itu, penelitian yang dilakukan oleh [8] mengenai deteksi bencana alam pada video citra satelit berdasarkan citra berbasis konten menggunakan metode CNN dengan arsitektur ResNet-50 menghasilkan sebuah sistem yang mampu mengolah citra satelit untuk memprediksi dan memperingatkan terjadinya angin topan dan kebakaran hutan, juga bisa mengolah citra satelit untuk mendeteksi wilayah terdampak banjir

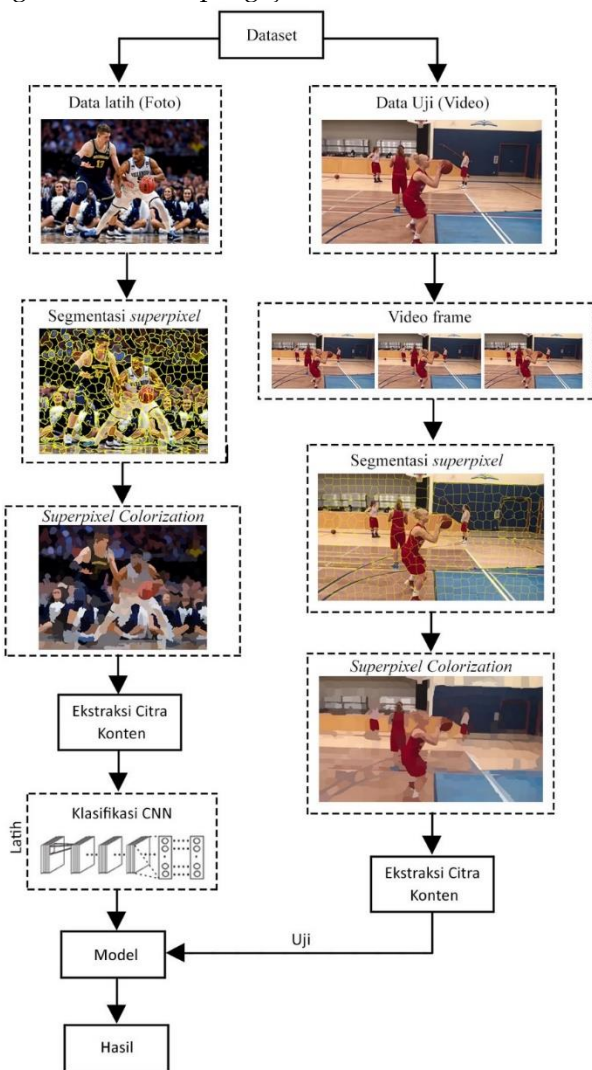
dan gempa bumi secara cepat. Prinsip *Rolling Averaging* (RA) digunakan saat pengujian video untuk memprediksi jenis bencana [8]. Proses klasifikasi citra juga dipengaruhi oleh cara mendapatkan fiturnya. Salah satu cara untuk mengambil fitur dari citra secara efektif dan mampu mengurangi kerumitan pemrosesan citra setelahnya adalah *superpixel* [9].

*Superpixel* mengelompokkan *pixel* yang memiliki kemiripan (homogen) menjadi satu daerah yang berisi informasi semantik dari suatu objek. *Superpixel* bisa digunakan untuk melakukan segmentasi, *contour closure*, pelacakan objek, rekonstruksi stereo 3D dan masih banyak lagi. *Superpixel* memiliki beberapa karakteristik, diantaranya setiap *superpixel* merupakan wilayah yang masing-masing citranya saling terhubung (*connectivity*), bentuk dan ukuran *superpixel* reguler di wilayah bukan fitur (*compactness*), *superpixel* mematuhi batasan segmentasi citra (*feature preservation*), kerapatan *superpixel* bersifat adaptif menyesuaikan konten dari citra (*content sensitivity*), mudah digunakan, dan memiliki performa serta akurasi yang baik [10]. Konten dari sebuah citra berupa warna, tekstur, corak dan bentuk juga bisa digunakan sebagai fitur untuk proses klasifikasi citra dengan baik karena menggunakan beberapa parameter dan telah dilakukan oleh [11], [12], [13], [14]. Penelitian yang dilakukan oleh [15] mengenai klasifikasi citra menggunakan segmentasi *superpixel* untuk deteksi kecurangan ujian pada video menghasilkan persentase rata-rata label terprediksi dengan benar adalah 90,064%. Selain itu, penelitian yang dilakukan oleh [16] mengenai klasifikasi penutup lahan fusi citra lidar dan citra optik udara dengan menerapkan *Superpixel Colorization* menghasilkan rata-rata akurasi sebesar 90,40%. Penelitian lain juga banyak menggunakan segmentasi *superpixel* untuk pengolahan citra [17], [18], [19], [20].

## 2. METODE DAN BAHAN

Pada penelitian ini dataset yang digunakan ada 2 jenis yaitu citra dan video. Dataset citra diambil dari *Sports Type Classifier* [21]. Dataset video diambil dari *Sport Video In The Wild* [22]. Olahraga yang digunakan ada 9 jenis, yaitu: *Baseball*, *Basketball*, *Boxing*, *Gymnastic*, *Hockey*, *Swimming*, *Tennis*, *Volleyball*, dan *Wrestling* yang merupakan olahraga irisan dari kedua dataset. Dataset citra yang digunakan sebanyak 495 citra untuk masing-masing jenis olahraga yang akan digunakan untuk pelatihan. Dataset video yang digunakan sebanyak 109 video untuk masing-masing jenis olahraga yang akan digunakan untuk pengujian. Resolusi citra berbasis konten dan video yang digunakan adalah 224 x 224 piksel [23].

Pada tahap pelatihan model, setiap citra dilakukan tahap *preprocessing*, yaitu dengan penerapan segmentasi *superpixel* dan *superpixel colorization*. Selanjutnya, hasil *superpixel colorization* yang telah dilakukan pada citra akan menjadi *input* pada metode CNN dengan arsitektur ResNet-50 untuk proses pelatihan sehingga menghasilkan model. Model klasifikasi citra olahraga dengan berbagai jumlah segmentasi *superpixel* yaitu 100, 200, 300, 400, 500, 600, 700, 800, 900, 1.000, 10.000, dan 25.000 dibentuk untuk dilakukan perbandingan hasil klasifikasi. Model yang menghasilkan tingkat *accuracy*, *precision*, dan *recall* yang paling baik akan digunakan untuk pengujian.



Gambar 1. Rancangan Klasifikasi Video Olahraga Berdasarkan Citra Berbasis Konten Menggunakan Segmentasi *Superpixel*.

Pada tahap pengujian, dilakukan proses *looping* untuk setiap *frame* pada video sesuai dengan prinsip RA. Setiap *frame* dilakukan proses segmentasi *superpixel*, *superpixel colorization*, dan *mean subtraction* untuk dinormalisasi sebelum akhirnya diujikan ke model. Setelah melakukan pengujian dengan menerapkan prinsip RA, keluarannya berupa label

nama olahraga. Rancangan proses klasifikasi video olahraga berdasarkan citra berbasis konten menggunakan segmentasi *superpixel* dapat dilihat pada Gambar 1.

### 2.1. Simple Linear Iterative Clustering (SLIC)

SLIC merupakan sebuah algoritma untuk menghasilkan *superpixel* yang terinspirasi dari algoritma klasterisasi *k-means*. SLIC dimulai dengan inialisasi pusat kluster  $C_i$  untuk menghasilkan *superpixel* pada ruang sampel  $S$ .  $S$  didapatkan dari Persamaan 1.

$$S = \sqrt{\frac{N}{K}} \quad (1)$$

dimana  $N$  adalah jumlah *pixel* dan  $K$  adalah jumlah *superpixel* yang diinginkan. Algoritma SLIC hanya mencari *pixel* serupa dari setiap pusat *cluster* ke *pixel*-nya untuk pengelompokan wilayah  $2S \times 2S$  sehingga bisa menghasilkan *superpixel* lebih cepat. Jarak  $D$  didapatkan dengan perhitungan *Euclidean Distance*:

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (2)$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (3)$$

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{s}\right)^2} \cdot t^2 \quad (4)$$

dimana  $l$ ,  $a$ , dan  $b$  adalah koordinat pada ruang warna CIELAB,  $x$  dan  $y$  adalah koordinat pada ruang  $xy$ ,  $d_c$  adalah jarak antara dua *pixel* pada ruang CIELAB,  $d_s$  adalah jarak antara dua *pixel* pada ruang  $xy$ , karena menggunakan ruang warna CIELAB,  $t$  berada pada rentang interval [1 40] [24]. Pada algoritma SLIC, waktu pemrosesan ikut dipengaruhi oleh kompleksitas  $O(N)$  dari jumlah segmentasi *superpixel* [25].

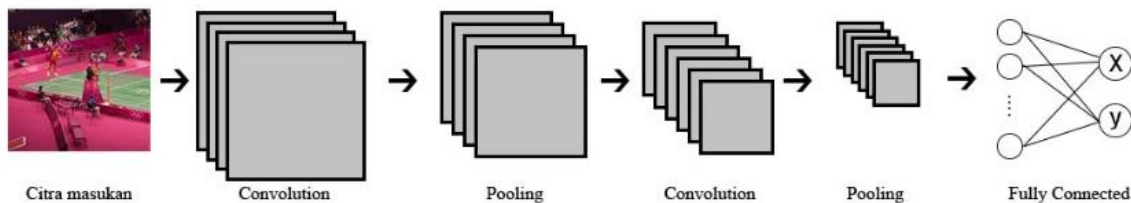
### 2.2. Superpixel Colorization (SC)

Pada tahap SC, dilakukan pewarnaan ulang untuk setiap piksel pada setiap segmentasi dengan menggunakan warna rata-ratanya.

### 2.3. Convolutional Neural Network (CNN)

CNN merupakan pengembangan dari *Multilayer Perceptron* (MLP) dan termasuk kedalam jenis *Deep Neural Network* karena memiliki tingkat kedalaman jaringan yang tinggi [26].

Arsitektur CNN dapat dilatih dan terdiri dari beberapa tahap. CNN terdiri dari satu lapis masukan, satu lapis keluaran dan beberapa lapis *hidden layer*.

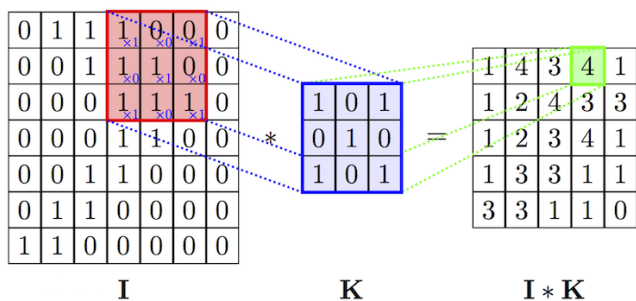


Gambar 2. Arsitektur CNN Secara Umum [27]

Struktur CNN berbeda-beda tergantung dari tipe arsitekturnya. Pada umumnya, struktur CNN terdiri dari *Convolutional Layer*, *Rectified Linear Units (ReLU) Layer*, *Pooling Layer*, dan *Fully Connected Layer*. Arsitektur CNN dapat dilihat pada Gambar 2.

### 2.3.1 Convolutional Layer

Lapisan ini terdiri dari *neuron* yang tersusun sedemikian rupa membentuk *filter* yang memiliki panjang dan tinggi. Lapisan ini bertujuan untuk mendapatkan fitur dari citra. *Filter* akan digeser keseluruhan bagian dari citra sambil melakukan operasi "dot" antara masukan dan nilai dari *filter* tersebut pada setiap pergeserannya untuk menghasilkan peta fitur. I adalah citra awal sebelum dikonvolusi, K adalah *kernel* atau *filter*, sedangkan  $I \times K$  adalah citra hasil konvolusi. Ilustrasi proses konvolusi dapat dilihat pada Gambar 3.



Gambar 3. Proses Konvolusi [27]

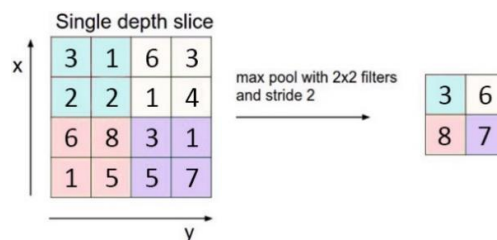
### 2.3.2 Rectified Linear Units Layer

Menerapkan fungsi aktivasi yang menormalisasi *pixel* dengan cara hanya meloloskan *pixel* dengan nilai positif, *pixel* yang bernilai negatif diubah menjadi 0. Fungsi aktivasinya dapat dilihat pada Persamaan 5

$$f(x) = \max(0, x) \quad (5)$$

### 2.3.3 Pooling Layer

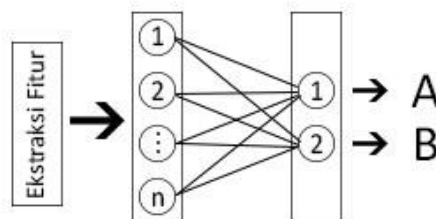
Pada lapisan ini dilakukan pengurangan dimensi peta fitur dan mempertajam fitur dengan cara membuat kernel kecil (biasanya berukuran 2x2 atau 3x3), kemudian kernel tersebut dijalankan pada citra dengan cara menggeser kernel tersebut untuk mengambil *neuron* nilai terbesar dari lingkup kernel, sehingga tercipta citra baru dengan dimensi yang lebih kecil. Ilustrasi proses *Pooling* dapat dilihat pada Gambar 4.



Gambar 4. Proses Mencari *Neuron* Bernilai Maksimal dengan *Pooling* [28]

### 2.3.4 Fully Connected Layer

Pada lapisan ini semua *neuron* aktivasi dari lapisan sebelumnya terhubung ke semua *neuron* di lapisan selanjutnya untuk dihitung bobotnya. Ilustrasi *Fully Connected Layer* dapat dilihat pada Gambar 5.

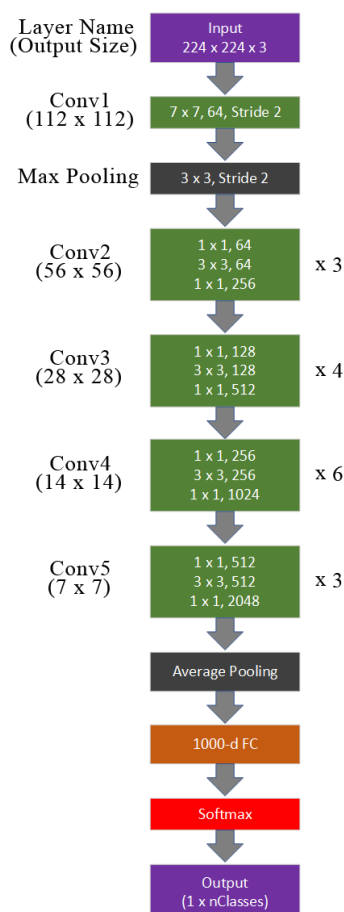


Gambar 5. *Fully Connected Layer* [28]

## 2.4. Arsitektur CNN ResNet-50

ResNet-50 adalah arsitektur dari CNN yang menggunakan konsep *shortcut connections*. *Shortcut connections* menggunakan kembali fitur yang dipakai sebagai masukan di lapisan sebelumnya terhadap keluaran dari lapisan itu sendiri untuk mengurangi hilangnya fitur-fitur penting saat melakukan konvolusi. ResNet-50 terdiri dari 5 tahap konvolusi, lalu dilanjutkan dengan *average pooling* dan *Fully Connected Layer*. *Input*-nya berupa citra berukuran 224 x 224 piksel dengan kedalaman 3. Pada tahap konvolusi pertama, dilakukan proses konvolusi dengan ukuran filter 7 x 7, kedalaman 64, dan langkah pergeseran *filter* (*stride*) sebanyak 2 piksel. *Output* dari konvolusi tahap pertama berukuran 112 x 112 piksel. Setelah itu dilakukan proses *Max Pooling* dengan ukuran *filter* 3 x 3, dan *stride*-nya 2 piksel. Pada tahap konvolusi kedua, dilakukan proses konvolusi dengan ukuran 1 x 1 dengan kedalaman 64, dilanjutkan konvolusi dengan ukuran 3 x 3 dengan kedalaman 64, lalu konvolusi 1 x 1 dengan

kedalaman 256. Proses konvolusi kedua ini dilakukan sebanyak tiga kali. *Output* dari konvolusi tahap kedua berukuran  $56 \times 56$  piksel. Pada tahap konvolusi ketiga, dilakukan proses konvolusi dengan ukuran  $1 \times 1$  dengan kedalaman 128, lalu dilanjutkan konvolusi dengan ukuran  $3 \times 3$  dengan kedalaman 128, kemudian konvolusi dengan ukuran  $1 \times 1$  dengan ukuran 512. Proses konvolusi ketiga ini dilakukan sebanyak 4 kali. *Output* dari konvolusi tahap ketiga berukuran  $28 \times 28$  piksel. Pada tahap konvolusi keempat, dilakukan proses konvolusi dengan ukuran  $1 \times 1$  dengan kedalaman 256, lalu dilanjutkan konvolusi dengan ukuran  $3 \times 3$  dengan kedalaman 256, kemudian konvolusi dengan ukuran  $1 \times 1$  dengan ukuran 1024. Proses konvolusi keempat ini dilakukan sebanyak 6 kali. *Output* dari konvolusi tahap keempat berukuran  $14 \times 14$  piksel. Pada tahap konvolusi kelima, dilakukan proses konvolusi dengan ukuran  $1 \times 1$  dengan kedalaman 512, lalu dilanjutkan konvolusi dengan ukuran  $3 \times 3$  dengan kedalaman 512, kemudian konvolusi dengan ukuran  $1 \times 1$  dengan ukuran 2048.



Gambar 6. Arsitektur ResNet-50 [29]

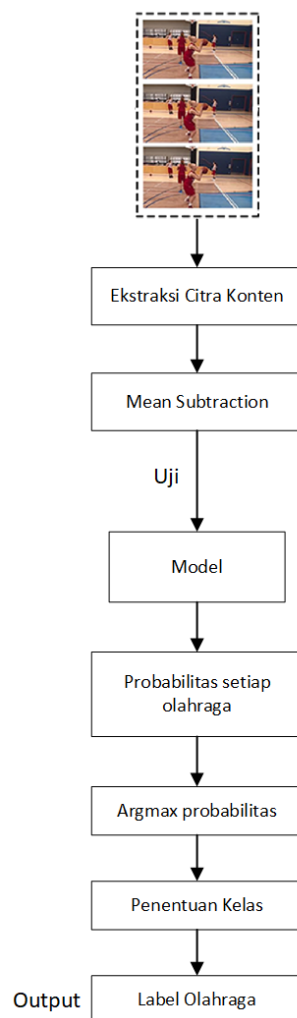
Proses konvolusi kelima ini dilakukan sebanyak 3 kali. *Output* dari konvolusi tahap kelima berukuran  $7 \times 7$  piksel. Setelah itu dilakukan proses *Average Pooling*, kemudian *Fully Connected* dengan 1000 *neuron*. Hasilnya dilakukan proses *softmax* untuk

menghasilkan *output* berupa kelas. Arsitektur ResNet-50 dapat dilihat pada Gambar 6.

## 2.5. Rolling Averaging (RA)

Prinsip RA digunakan untuk menentukan label jenis olahraga sebagai hasil *output* dari klasifikasi video. Dilakukan proses *looping* untuk setiap *frame* dari video untuk diujikan ke model CNN. Hasil prediksi setiap *frame* pada video akan dihitung probabilitas untuk setiap jenis olahraganya. Kemudian diambil probabilitas terbesar pada kandidat probabilitas untuk jenis olahraga ke-*i* di setiap *frame* dan mewakili sebagai probabilitas tiap jenis olahraga. Hasil pengujiannya berupa label dari kelas jenis olahraga yang memiliki probabilitas terbesar.

Frame-frame video telah dilakukan SS dan SC



Gambar 7. Skema Rolling Averaging

Proses menghitung probabilitas dan penentuan kelasnya dapat dilihat pada Persamaan 6 dan Persamaan 7.





Gambar 8. Citra Hasil *Superpixel Colorization* dengan Berbagai Jumlah *Superpixel*

$$c = \arg \max(P_{s_1}, P_{s_2}, \dots, P_{s_n}) \quad (7)$$

dimana  $P_{s_i}$  adalah probabilitas maksimum olahraga ke- $i$  dideteksi sebagai suatu olahraga diantara semua *frame* dalam video,  $f_k$  adalah *frame* ke- $k$ ,  $k$  adalah jumlah *frame* pada video,  $c \in C$  dimana  $C = \{C_1, C_2, \dots, C_n\}$  merupakan kelas-kelas olahraga,  $n$  adalah banyaknya jenis olahraga. Skema perancangan prinsip RA dapat dilihat pada Gambar 7.

### 3. HASIL DAN PEMBAHASAN

Citra hasil segmentasi *superpixel* dan *superpixel colorization* pada berbagai jumlah segmentasi dapat dilihat pada Gambar 8.

Hasil pengujian model klasifikasi citra olahraga

berdasarkan jumlah segmentasi *superpixel* dapat dilihat pada Tabel 1, Tabel 2, Tabel 3 dan Tabel 4, bisa dilihat bahwa secara keseluruhan tingkat *accuracy*, *precision*, dan *recall* dari model mengalami peningkatan.

Berdasarkan data hasil pengujian berdasarkan jumlah segmentasi *superpixel* mulai dari 100 sampai 25.000 pada Tabel 1, Tabel 2, Tabel 3 dan Tabel 4, jumlah segmentasi yang memiliki hasil *accuracy*, *precision*, dan *recall* yang paling baik adalah 25.000. Sesuai dengan hasil tersebut, maka selanjutnya dilakukan pengujian video menggunakan model klasifikasi olahraga dengan jumlah segmentasi *superpixel* sebanyak 25.000.

Grafik rata-rata hasil pengujian model menggunakan jumlah segmentasi *superpixel* secara keseluruhan dapat dilihat pada Gambar 9.

Tabel 1. Perbandingan Hasil untuk Jumlah Segmentasi *Superpixel* 100 s.d. 300

No.	Olahraga	100 sp			200 sp			300 sp		
		Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
1	Baseball	0,95	0,75	0,81	0,94	0,76	0,70	0,96	0,80	0,83
2	Basketball	0,85	0,39	0,60	0,88	0,47	0,55	0,88	0,46	0,68
3	Boxing	0,93	0,67	0,67	0,94	0,67	0,83	0,93	0,64	0,79
4	Gymnastic	0,92	0,84	0,37	0,91	0,73	0,35	0,93	0,83	0,52
5	Hockey	0,94	0,70	0,94	0,98	0,84	0,96	0,96	0,74	0,95
6	Swimming	0,98	0,92	0,94	0,99	0,98	0,96	0,98	0,92	0,91
7	Tennis	0,93	0,64	0,60	0,93	0,66	0,70	0,93	0,72	0,72
8	Volleyball	0,92	0,64	0,60	0,93	0,66	0,77	0,92	0,67	0,55
9	Wrestling	0,91	0,62	0,48	0,92	0,64	0,55	0,91	0,64	0,34
<b>AVERAGE</b>		<b>0,92</b>	<b>0,68</b>	<b>0,67</b>	<b>0,94</b>	<b>0,71</b>	<b>0,71</b>	<b>0,93</b>	<b>0,71</b>	<b>0,70</b>

Tabel 2. Perbandingan Hasil untuk Jumlah Segmentasi *Superpixel* 400 s.d. 600

No.	Olahraga	400 sp			500 sp			600 sp		
		Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
1	Baseball	0,95	0,86	0,68	0,95	0,74	0,80	0,95	0,84	0,70
2	Basketball	0,90	0,54	0,53	0,88	0,47	0,74	0,91	0,61	0,56
3	Boxing	0,95	0,74	0,85	0,94	0,72	0,76	0,95	0,75	0,84
4	Gymnastic	0,93	0,81	0,44	0,93	0,85	0,42	0,93	0,86	0,40
5	Hockey	0,97	0,81	0,95	0,96	0,76	0,94	0,97	0,84	0,96
6	Swimming	0,99	0,96	0,98	0,98	0,95	0,90	0,99	0,95	0,95
7	Tennis	0,93	0,68	0,76	0,94	0,79	0,67	0,94	0,74	0,70
8	Volleyball	0,92	0,63	0,76	0,92	0,63	0,72	0,90	0,53	0,87
9	Wrestling	0,92	0,65	0,65	0,92	0,69	0,42	0,93	0,70	0,65
<b>AVERAGE</b>		<b>0,94</b>	<b>0,74</b>	<b>0,73</b>	<b>0,94</b>	<b>0,73</b>	<b>0,71</b>	<b>0,94</b>	<b>0,76</b>	<b>0,74</b>

Tabel 3. Perbandingan Hasil untuk Jumlah Segmentasi *Superpixel* 700 s.d. 900

No.	Olahraga	700 sp			800 sp			900 sp		
		Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
1	Baseball	0,95	0,79	0,81	0,94	0,81	0,67	0,96	0,86	0,80
2	Basketball	0,90	0,53	0,61	0,90	0,54	0,66	0,90	0,54	0,78
3	Boxing	0,95	0,73	0,84	0,94	0,68	0,85	0,94	0,72	0,80
4	Gymnastic	0,94	0,78	0,59	0,93	0,80	0,47	0,94	0,82	0,57
5	Hockey	0,96	0,78	0,96	0,97	0,85	0,94	0,97	0,82	0,96
6	Swimming	0,98	0,96	0,88	0,99	0,96	0,96	0,98	0,96	0,88
7	Tennis	0,94	0,75	0,72	0,94	0,72	0,75	0,94	0,74	0,74
8	Volleyball	0,93	0,69	0,65	0,93	0,65	0,85	0,93	0,71	0,66
9	Wrestling	0,91	0,62	0,52	0,94	0,82	0,53	0,92	0,65	0,50
<b>AVERAGE</b>		<b>0,94</b>	<b>0,74</b>	<b>0,73</b>	<b>0,94</b>	<b>0,76</b>	<b>0,74</b>	<b>0,94</b>	<b>0,76</b>	<b>0,74</b>

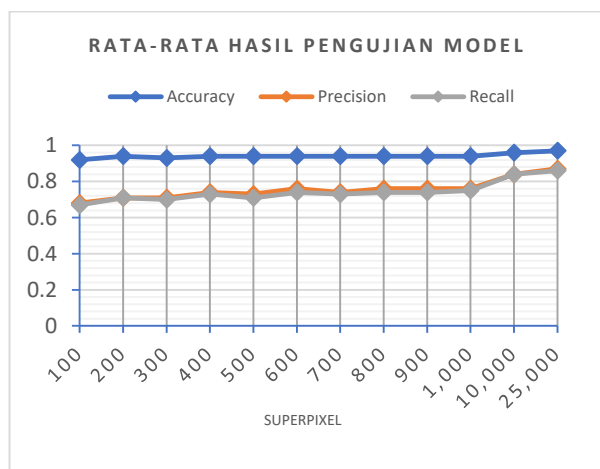


Tabel 4. Perbandingan Hasil untuk Jumlah Segmentasi *Superpixel* di atas 900

No.	Olahraga	1.000 sp			10.000 sp			25.000 sp		
		Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
1	Baseball	0,95	0,84	0,74	0,97	0,87	0,86	0,98	0,90	0,94
2	Basketball	0,91	0,58	0,64	0,94	0,71	0,85	0,96	0,78	0,87
3	Boxing	0,94	0,71	0,84	0,96	0,82	0,88	0,96	0,82	0,86
4	Gymnastic	0,92	0,83	0,40	0,95	0,85	0,69	0,96	0,88	0,73
5	Hockey	0,98	0,88	0,94	0,98	0,90	0,98	0,99	0,94	0,97
6	Swimming	0,98	0,93	0,93	0,99	0,97	0,93	0,99	0,98	0,93
7	Tennis	0,94	0,70	0,80	0,97	0,88	0,82	0,97	0,90	0,84
8	Volleyball	0,93	0,64	0,83	0,96	0,81	0,85	0,96	0,82	0,90
9	Wrestling	0,93	0,76	0,62	0,94	0,76	0,68	0,95	0,80	0,75
	<b>AVERAGE</b>	<b>0,94</b>	<b>0,76</b>	<b>0,75</b>	<b>0,96</b>	<b>0,84</b>	<b>0,84</b>	<b>0,97</b>	<b>0,87</b>	<b>0,86</b>

Tabel 5. *Confusion Matrix* Pengujian Video Olahraga dengan 25.000 sp

		Predicted								
		Baseball	Basketball	Boxing	Gymnastic	Hockey	Swimming	Tennis	Volleyball	Wrestling
Ground Truth	Baseball	66	5	6	7	4	1	4	11	5
	Basketball	5	20	13	26	2	4	10	25	4
	Boxing	1	0	69	12	4	0	2	11	10
	Gymnastic	0	1	5	101	0	0	1	0	1
	Hockey	3	0	2	6	78	1	0	18	1
	Swimming	8	0	7	7	1	72	0	12	2
	Tennis	11	3	5	5	0	1	75	9	0
	Volleyball	1	2	8	21	3	1	4	69	0
	Wrestling	0	7	5	23	10	0	2	14	48



Gambar 9. Grafik Rata-rata Hasil Pengujian Model

*Confusion Matrix* hasil klasifikasi video olahraga menggunakan model klasifikasi dengan jumlah segmentasi *superpixel* sebanyak 25.000 dapat dilihat pada Tabel 5. Berdasarkan data yang ditampilkan pada Tabel 5, 3 jenis video olahraga yang paling banyak diklasifikasi dengan benar dari 109 video yang diujikan adalah *gymnastic* sebanyak 101 video, *hockey* sebanyak 78 video, dan *tennis* sebanyak 75 video. Sedangkan, 3 jenis video olahraga yang paling sedikit diklasifikasi dengan benar dari 109 video yang diujikan adalah *basketball* sebanyak 20 video, *wrestling* sebanyak 48 video, dan *baseball* sebanyak 66 video. Pada pengujian video jenis olahraga *wrestling*, klasifikasi masih cukup baik dengan didominasi sebanyak 48 video dengan benar, namun terdapat 23

video yang diklasifikasi sebagai jenis olahraga *gymnastic*. Sedangkan, pada pengujian video jenis olahraga *basketball* masih diklasifikasi dengan jenis olahraga yang belum tepat, hal ini dapat dilihat pada Tabel 5.

Video olahraga *basketball* sendiri hanya diklasifikasi dengan benar sebanyak 20 video sehingga hasil klasifikasi video *basketball* kurang baik karena didominasi oleh olahraga *gymnastic* dan *volleyball*. Tingkat *accuracy*, *precision*, dan *recall* dari pengujian untuk masing-masing olahraga dapat dilihat pada Tabel 6.

Tabel 6. Hasil Pengujian Video Setiap Jenis Olahraga

No.	Olahraga	Acc	Prec	Rec	TP	FN	FP	TN
1	Baseball	0,93	0,69	0,60	66	43	29	843
2	Basketball	0,89	0,53	0,18	20	89	18	854
3	Boxing	0,91	0,58	0,63	69	40	51	821
4	Gymnastic	0,88	0,48	0,93	101	8	107	765
5	Hockey	0,94	0,76	0,72	78	31	24	848
6	Swimming	0,95	0,90	0,66	72	37	8	864
7	Tennis	0,94	0,76	0,69	75	34	23	849
8	Volleyball	0,86	0,41	0,63	69	40	100	772
9	Wrestling	0,91	0,68	0,44	48	61	23	849
	<b>AVERAGE</b>	<b>0,91</b>	<b>0,64</b>	<b>0,61</b>				

Berdasarkan data yang diperoleh dari setiap jenis video olahraga pada Tabel 6, didapatkan hasil yang terbaik dari olahraga *swimming* dengan *accuracy* 0,95; *precision* 0,90; dan *recall* 0,66; kemudian diikuti oleh olahraga *hockey* dengan *accuracy* 0,94; *precision* 0,76; dan *recall* 0,72; lalu olahraga *tennis* dengan *accuracy* 0,94; *precision* 0,76; dan *recall* 0,69. Tingkat olahraga



yang memiliki hasil yang kurang baik, yaitu *basketball* dengan *accuracy* 0,89; *precision* 0,53; dan *recall* 0,18; kemudian olahraga *volleyball* dengan *accuracy* 0,86; *precision* 0,41; dan *recall* 0,63; lalu olahraga *wrestling* dengan *accuracy* 0,91; *precision* 0,68; dan *recall* 0,44. Pada dataset video yang sama, telah dilakukan klasifikasi video dengan menggunakan metode CNN yang dikombinasikan dengan sistem *keyframe extractor* dan didapatkan hasil rata-rata *accuracy* sebesar 0,85 [30]. Sedangkan, pada penelitian ini dilakukan klasifikasi video dengan menggunakan metode segmentasi *superpixel* yang memberikan hasil yang baik pula dengan rata-rata *accuracy* 0,91; *precision* 0,64; dan *recall* 0,61. Walaupun dataset yang digunakan pada pelatihan dan pengujian sama-sama berjenis olahraga, tetapi dataset yang digunakan berasal dari sumber yang berbeda. Beberapa olahraga seperti *basketball*, *volleyball*, dan *wrestling* tidak terklasifikasi dengan baik karena konten citra yang dipelajari dari dataset pelatihan memiliki karakteristik yang berbeda dibanding dengan konten frame video olahraga yang diujikan atau juga disebabkan oleh warna konten dan warna di luar konten yang dianggap mirip sehingga tidak mencerminkan jenis olahraga tersebut.

#### 4. KESIMPULAN

Berdasarkan pengujian model klasifikasi pada jumlah segmentasi *superpixel* yang digunakan, maka dapat disimpulkan bahwa segmentasi *superpixel* bisa diterapkan untuk mengklasifikasikan video olahraga berdasarkan citra berbasis konten. Kemudian, berdasarkan Tabel 1 s.d. Tabel 4 pembentukan model terbaik menggunakan jumlah segmentasi *superpixel* sebanyak 25.000. Sesuai dengan data pada Tabel 6, hasil dari pengujian video olahraga yaitu rata-rata *accuracy* 0,91; *precision* 0,64; dan *recall* 0,61. Berdasarkan Tabel 5, olahraga yang diklasifikasi paling baik adalah *gymnastic* dengan jumlah video yang berhasil diklasifikasi dengan benar sebanyak 101 video dari total 109 video. Sedangkan hasil klasifikasi video olahraga yang paling rendah adalah *basketball* dengan jumlah video yang berhasil diklasifikasi dengan benar hanya 20 video dari total 109 video. Klasifikasi pada video olahraga *basketball* masih kurang baik dengan tingkat *recall* 0,18 karena didominasi oleh kemiripan warna pada arena jenis olahraga *gymnastic* dan *volleyball*. Saran untuk penelitian selanjutnya adalah melakukan klasifikasi video olahraga berdasarkan citra berbasis konten menggunakan segmentasi *superpixel* dengan optimasi ataupun menggunakan arsitektur CNN selain ResNet-50 berdasarkan citra berbasis konten.

#### DAFTAR PUSTAKA

- [1] Y. Hari, "Latihan Fisik dan Kekebalan Tubuh," *Medikora*, vol. IV, no. 1, pp. 47–65, 2008.
- [2] N. R. Kasat and S. D. Thepade, "Novel Content Based Image Classification Method Using LBG Vector Quantization Method with Bayes and Lazy Family Data Mining Classifiers," 2016, doi: 10.1016/j.procs.2016.03.062.
- [3] M. Ramesh and K. Mahesh, "A Performance Analysis of Pre-trained Neural Network and Design of CNN for Sports Video Classification," Sep. 2020, pp. 0213–0216, doi: 10.1109/icccsp48568.2020.9182113.
- [4] D. S. Sachan, U. Tekwani, and A. Sethi, "Sports video classification from multimodal information using deep neural networks," 2013.
- [5] J. Hanna, F. Patlar, A. Akbulut, E. Mendi, and C. Bayrak, "HMM based classification of sports videos using color feature," 2012, doi: 10.1109/IS.2012.6335247.
- [6] S. Uma Maheswari and R. Ramakrishnan, "Sports Video Classification Using Multi Scale Framework and Nearest Neighbor Classifier," *Indian J. Sci. Technol.*, 2015, doi: 10.17485/ijst/2015/v8i6/61067.
- [7] P. Campr, M. Herbig, J. Vanek, and J. Psutka, "Sports video classification in continuous TV broadcasts," 2014, doi: 10.1109/ICOSP.2014.7015083.
- [8] G. Kumar, "Average Rolling based Real-time Calamity Detection using Deep Learning," India, 2019. [Online]. Available: <https://towardsdatascience.com/average-rolling-based-real-time-calamity-detection-using-deep-learning-ae51a2ffd8d2>.
- [9] H. Zou, X. Qin, S. Zhou, and K. Ji, "A likelihood-based SLIC superpixel algorithm for SAR images using generalized gamma distribution," *Sensors (Switzerland)*, 2016, doi: 10.3390/s16071107.
- [10] Y. J. Liu, M. Yu, B. J. Li, and Y. He, "Intrinsic Manifold SLIC: A Simple and Efficient Method for Computing Content-Sensitive Superpixels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 656–666, 2018, doi: 10.1109/TPAMI.2017.2686857.
- [11] R. Bibi, Z. Mehmood, R. M. Yousaf, T. Saba, M. Sardaraz, and A. Rehman, "Query-by-visual-search: multimodal framework for content-based image retrieval," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2020, doi: 10.1007/s12652-020-01923-1.
- [12] A. H. Rangkuti, Z. E. Rasjid, and D. J. Santoso, "Batik Image Classification Using Treeval and Treefit as Decision Tree Function in Optimizing Content Based Batik Image Retrieval," 2015, doi: 10.1016/j.procs.2015.07.551.
- [13] R. Das, S. De, S. Bhattacharyya, J. Platos, V. Snasel, and A. E. Hassaniien, "Data

- Augmentation and Feature Fusion for Melanoma Detection with Content Based Image Classification," 2020, doi: 10.1007/978-3-030-14118-9\_70.
- [14] S. R. Shinde, S. Sabale, S. Kulkarni, and D. Bhatia, "Experiments on content based image classification using Color feature extraction," 2015, doi: 10.1109/ICCICT.2015.7045737.
- [15] Yohannes, V. Ayumi, and M. I. Fanany, "Multimodal decomposable models by superpixel segmentation and point-in-time cheating detection," *2016 Int. Conf. Adv. Comput. Sci. Inf. Syst.*, pp. 391–396, 2016, doi: 10.1109/ICACIS.2016.7872729.
- [16] N. Rachmat, "Klasifikasi penutup lahan fusi citra lidar dan citra optik udara menggunakan superpixel dan maximum likelihood," Universitas Indonesia, 2017.
- [17] W. Sultani, S. Mokhtari, and H. B. Yun, "Automatic Pavement Object Detection Using Superpixel Segmentation Combined with Conditional Random Field," *IEEE Trans. Intell. Transp. Syst.*, 2018, doi: 10.1109/TITS.2017.2728680.
- [18] P. Ebenezer, H. Nunoo-Mensah, R. Junior, and J. Raissa, "Brain Tumor Segmentation using SLIC Superpixels and Optimized Thresholding Algorithm," *Int. J. Comput. Appl.*, 2018, doi: 10.5120/ijca2018917915.
- [19] Y. Yohannes, "Deteksi Teks Secara Otomatis Pada Natural Image Berbasis Superpixel Menggunakan Maximally Stable Extremal Regions dan Stroke Width Transform," *J. Tek. Inform. dan Sist. Inf.*, vol. 3, no. 2, pp. 332–343, 2017, doi: 10.28932/jutisi.v3i2.630.
- [20] Y. Yohannes, N. Rachmat, and C. O. Saputra, "Penggunaan Fitur HOG Berbasis Superpixel Untuk Klasifikasi Jenis Jamur Dengan Metode SVM," *Jusikom J. Sist. Komput. Musirawas*, vol. 6, no. 1, pp. 23–31, 2021, doi: 10.32767/jusikom.v6i1.1180.
- [21] S. L. Jurj, "Sports Type Classifier," 2019. <https://github.com/jurjsorinliviu/Sports-Type-Classifier> (accessed Aug. 20, 2020).
- [22] S. M. Safdarnejad, X. Liu, L. Udpa, B. Andrus, J. Wood, and D. Craven, "Sports Videos in the Wild (SVW): A video dataset for sports analysis," *2015 11th IEEE Int. Conf. Work. Autom. Face Gesture Recognition, FG 2015*, 2015, doi: 10.1109/FG.2015.7163105.
- [23] A. Kwasigroch, A. Mikołajczyk, and M. Grochowski, "Deep neural networks approach to skin lesions classification - A comparative analysis," 2017, doi: 10.1109/MMAR.2017.8046978.
- [24] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2281, 2012, doi: 10.1109/TPAMI.2012.120.
- [25] A. L. Radhakrishna Achanta, Appu Shaji, Kevin Smith and S. S. Pascal Fua, "SLIC Superpixels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2010, doi: 10.1109/TPAMI.2012.120.
- [26] T. Nurhikmat and T. Purwaningsih, "Implementasi Deep Learning Untuk Image Classification Menggunakan Algoritma Convolutional Neural Network (Cnn) Pada Citra Wayang Golek," *ResearchGate*, 2018, doi: 10.13140/RG.2.2.10880.53768.
- [27] J. P. Mueller and L. Massaron, *Deep Learning for Dummies*. 2019.
- [28] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, and V. Zocca, *Python Deep Learning: Exploring Deep Learning Techniques and Neural Network Architectures with PyTorch, Keras and TensorFlow*. 2018.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016, doi: 10.1109/CVPR.2016.90.
- [30] R. F. Rachmadi, K. Uchimura, and G. Koutaki, "Video classification using compacted dataset based on selected keyframe," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, pp. 873–878, 2017, doi: 10.1109/TENCON.2016.7848130.