

GETARI: DATASET UNTUK KLASIFIKASI GERAKAN DASAR TARI BALI PEREMPUAN

I Putu Putra Budha Lantara¹, I Putu Dwi Payana², Gede Ariel Septian Pratama³, Wayan Evan Ada Munayana⁴, Kadek Sri Nopiani⁵, I Nyoman Rudy Hendrawan⁶

Sistem Informasi, Institut Teknologi dan Bisnis STIKOM Bali

email: budhalantara@gmail.com¹, ini.dwiii@gmail.com², arielseptian1299@gmail.com³, yanevan03@gmail.com⁴, novi.resni@gmail.com⁵, rudyhendrawan@stikom-bali.ac.id⁶

Abstrak

Tari Bali adalah salah satu daya tarik kultural yang selalu dilestarikan oleh masyarakat Bali dari dulu hingga sekarang. Hingga saat ini banyak upaya telah dilakukan oleh masyarakat untuk mengabadikan karya seni Tari Bali ke dalam media digital. Salah satu pendekatan yang dapat digunakan dalam rangka preservasi karya seni Tari Bali yaitu dengan pendekatan teknologi machine learning. Pada penelitian ini dilakukan pembuatan dataset gerakan dasar Tari Bali yang merupakan kelanjutan dari penelitian sebelumnya yaitu, enam gerakan dasar Tari Bali yakni, *ngelung*, *ngeseh*, *tapak sirangpada*, *ngeed*, *ngelo*, dan *ngumbang*, yang dinamakan GETARI. GETARI kemudian diklasifikasikan dengan menggunakan model *pre-trained* VGG-LSTM dan I3D. Berdasarkan hasil *training* pada model VGG-LSTM didapatkan nilai *validation loss* sebesar 1,01 dan akurasi sebesar 0,57, sedangkan I3D memperoleh nilai *validation loss* sebesar 0,03 dan akurasi sebesar 0,97. Perubahan strategi training menurunkan *validation loss* VGG-LSTM menjadi sebesar 0,18 dengan akurasi mencapai 0,94. Selain itu dilakukan juga pengukuran terhadap metrik klasifikasi lainnya seperti *precision*, *recall*, dan F1. Secara keseluruhan, kinerja pada data *test* memperlihatkan bahwa model I3D tetap menjadi yang terbaik diantara keduanya, dengan nilai akurasi sebesar 0,97, *precision* sebesar 0,98, *recall* sebesar 0,98, dan F1 juga sebesar 0,98. Penelitian ini merupakan salah tahap awal dari pengembangan dataset gerakan dasar Tari Bali dan juga pengembangan model *machine learning* untuk mengklasifikasikan gerakan dasar Tari Bali. Model yang dikembangkan dapat dijadikan sebagai acuan dalam pengembangan model *machine learning* lainnya.

Kata kunci: Tari Bali, klasifikasi, GETARI, VGG-LSTM, I3D

Abstract

Balinese traditional is one of the cultural traditions that people have retained throughout history. The community has made several efforts to preserve Balinese dance artworks in digital formats. In the context of conserving Balinese dance art, machine learning technology is one approach that may be utilized. In this study, a dataset of the basic movements of Balinese dance (GETARI) was created, namely *ngelung*, *ngeseh*, *tapak sirangpada*, *ngeed*, *ngelo*, and *ngumbang*. The basic movement dataset was then classified using the VGG-LSTM and I3D *pre-trained* models. Experiment result shows that VGG-LSTM *validation loss* value is 1.01 and the accuracy is 0.57, whereas I3D attained *validation loss* of 0.03 and an accuracy of 0.97. Changes to the training strategy reduce the VGG-LSTM *validation loss* to 0.18 with an accuracy of 0.94. In addition, several classification measures such as *precision*, *recall*, and F1 were measured. Overall, the performance on the data test demonstrates that the I3D model is superior than the VGG-LSTM, with an accuracy value of 0.97, *precision* of 0.98, *recall* of 0.98, and F1 value of 0.98. This project is one of the earliest phases of generating a dataset of fundamental Balinese dancing motions and a machine learning model for classifying the dataset. This model may be used to build other machine learning models.

Keywords : Balinese Traditional Dance, GETARI, classification, VGG-LSTM, I3D

Diterima Redaksi: 27-09-2022 | Selesai Revisi: 09-12-2022 | Diterbitkan Online: 27-12-2022

DOI: <https://doi.org/10.23887/janapati.v11i3.52598>

PENDAHULUAN

Bali adalah salah satu pulau di Indonesia yang memiliki daya tarik kultural dan alam yang menjadi andalan pada sektor pariwisata. Tari Bali adalah salah satu daya tarik kultural yang selalu dilestarikan oleh masyarakat Bali dari dulu hingga sekarang. Keragaman jenis tarian yang ada di Bali pada umumnya dikelompokkan berdasarkan kegunaannya yaitu, untuk upacara keagamaan hingga rekreasional.

Hingga saat ini banyak upaya telah dilakukan oleh masyarakat untuk mengabadikan karya seni Tari Bali ke dalam media digital misalnya ke dalam *platform* media sosial berbagi gambar atau video seperti YouTube, TikTok, Instagram, hingga Facebook, ataupun yang hanya sebatas diunggah ke dalam media penyimpanan *cloud* seperti Google Drive, OneDrive, hingga Dropbox. Berbagai studi juga sudah dilakukan dalam rangka preservasi karya seni Tari Bali [1]–[3].

Pendekatan lainnya dapat digunakan dalam rangka preservasi karya seni Tari Bali yaitu dengan pendekatan teknologi pembelajaran mesin (*machine learning*). Pendekatan dengan teknologi *machine learning* merupakan pendekatan preservasi yang progresif, hal ini dikarenakan informasi berupa data video yang disimpan sedemikian rupa sehingga dapat digunakan dalam proses *machine learning*. Proses *machine learning* ini secara umum melingkupi proses yang di dalamnya terdapat proses pengolahan data yang intensif dari data mentah hingga data yang siap digunakan, sampai dengan proses pemodelan dan evaluasi model *machine learning*. Disamping itu, model *machine learning* yang dikembangkan dapat diimplementasikan ke dalam berbagai bentuk perangkat lunak.

Banyak studi yang telah dilakukan berdasarkan pendekatan ini namun tidak terikat pada Tari Bali saja. Tarian tradisional dari propinsi Aceh, Likok Pulo, dimodelkan dengan teknik Hidden Markov Model (HMM). Keseluruhan gerakan tari dimodelkan dengan teknik tersebut dengan cara merepresentasikan gestur tari sebagai *hidden discrete states* dan gerakan tarian sebagai fase [4]. Pendekatan dengan HMM ini digunakan juga pada saat merepresentasikan pose pada gerakan tarian dan pada saat proses pembelajaran dan rekognisi gerakan tarian [5], [6]. Beberapa peneliti menggunakan HMM dalam studinya [7]–[9], hal ini menunjukkan HMM merupakan teknik yang populer dalam memodelkan rekognisi gerakan karena kemampuan model ini dalam

merepresentasikan gestur pada suatu aktivitas manusia.

Secara umum gerakan tarian merupakan media ekspresi manusia yang menggunakan ritme gerakan tubuh secara estetik [10]. Pada kasus Tari Bali, studi awal menunjukkan bahwa pendekatan yang serupa dengan HMM digunakan untuk merekognisi gerakan Tari Bali. Studi awal ini [10]–[12] menggunakan *skeleton descriptor* dalam merepresentasikan titik tubuh penari dan analisis kluster sebagai teknik dalam mengelompokkan titik tubuh yang memiliki tingkat kedekatan tertentu.

Sejauh penelusuran studi pustaka yang telah dilakukan oleh penulis, hanya ada sedikit studi yang membahas mengenai topik ini yaitu studi oleh Kesiman, dkk [13]. Pada studi tersebut [13], dilakukan penelitian untuk mengungkap karakteristik Tari Bali yang diciptakan oleh seniman maestro dengan cara menganalisis pola *silhouette sequence*. Analisis ini dilakukan dengan menggunakan teknik ekstraksi fitur *Histogram of Gradient* (HoG) dan *Scale Invariant Features Transform* (SIFT) *Descriptor*. Adapun penelitian oleh [14]–[16] juga membahas mengenai pemodelan gerakan dasar tarian tradisional dari propinsi Jawa Timur, yaitu Tari Remo. Peneliti memodelkan gerakan tarian tersebut dengan menggunakan model *deep learning Generative Long Short-Term Memory* [14], dan membahas mengenai strategi training pada model tersebut [15], dan membandingkannya dengan model *Gated Recurrent Unit* (GRU).

Berdasarkan telurusan studi literatur, belum ada penelitian yang membahas tentang klasifikasi gerakan dasar Tari Bali, terutama klasifikasi gerakan dasar dengan pendekatan teknik ataupun metode lainnya seperti yang dijelaskan pada penelitian terdahulu di atas. Oleh karena itu, pada penelitian ini dilakukan pembuatan dataset gerakan dasar Tari Bali yang merupakan kelanjutan dari penelitian sebelumnya [17], di mana pada penelitian tersebut dilakukan pembuatan dataset tujuh gerakan dasar Tari Bali yaitu, *agem* kanan, *agem* kiri, *ulap-ulap*, *ngegol*, *seledet*, *nyalud*, dan *nyeregseg* [18], [19]. Dataset gerakan dasar tersebut kemudian diklasifikasikan dengan menggunakan model pre-trained VGG [20] dan I3D [21]. Sedangkan pada penelitian ini, penulis menambah enam gerakan tambahan yakni, *ngelung*, *ngeseh*, *tapak sirangpada*, *ngeed*, *ngelo*, dan *ngumbang*, dengan teknik klasifikasi dengan model pre-trained yang sama. Gerakan-gerakan tari tersebut adalah beberapa gerakan dasar bagi penari perempuan, di mana masih ada beberapa gerakan yang belum dapat

dibahas pada penelitian ini. Penulis belum mampu membahas gerakan-gerakan dasar Tari Bali lainnya karena keterbatasan sumber daya. Penjelasan mengenai teknik pengambilan data dan model *pre-trained* dijelaskan pada bagian selanjutnya.

METODE

Secara garis besar penelitian ini memiliki 5 tahapan utama yaitu pengumpulan *dataset*, *preprocessing* data, pemodelan, *training* model, dan pengujian. Gambar 2 menunjukkan diagram alir tahapan-tahapan penelitian.

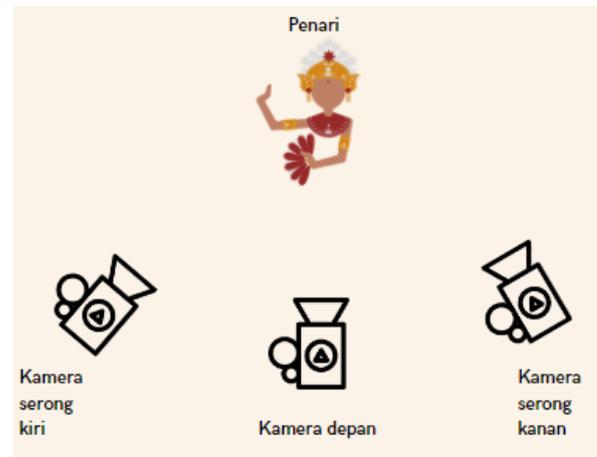
1. Dataset

Dataset dalam penelitian ini adalah video gerakan dasar Tari Bali yang diperagakan khusus oleh penari perempuan saja. Teknik pengambilan data yang dilakukan terinspirasi dari penelitian oleh Tsuchida [22]. Pada penelitian tersebut dilakukan pembuatan dataset mengenai tarian modern dengan berbagai jenis genre, banyak penari, dan banyak kamera.

Pada penelitian ini, video gerakan tari direkam menggunakan kamera *smartphone* dengan spesifikasi yang mampu melakukan perekaman video dengan resolusi 480 hingga 720 *progressive scan*. Pengambilan video menggunakan tiga buah *smartphone* dengan konfigurasi penempatan yaitu di depan, serong kiri dan serong kanan dari penari seperti yang diilustrasikan pada Gambar 1. Gerakan tari diperagakan lima orang penari dari Unit Kegiatan Mahasiswa (UKM) Tari Institut Teknologi dan Bisnis STIKOM Bali. Pengambilan video dilakukan di Laboratorium Seni dan Budaya dalam dua sesi pengambilan video. Tujuan dari dilakukannya pengambilan video dalam dua sesi adalah untuk mendapatkan video dengan objek latar belakang yang berbeda.

Setiap penari memeragakan enam gerakan dasar Tari Bali tersebut. Setiap gerakan diulang sebanyak lima kali di mana terdapat jeda selama 2 sampai dengan 3 detik pada setiap pengulangan. Sehingga jumlah keseluruhan video sebanyak 5 penari × 6 gerakan × 3 kamera = 90 video. Informasi

keseluruhan mengenai pengambilan video dapat dilihat pada Gambar 3.

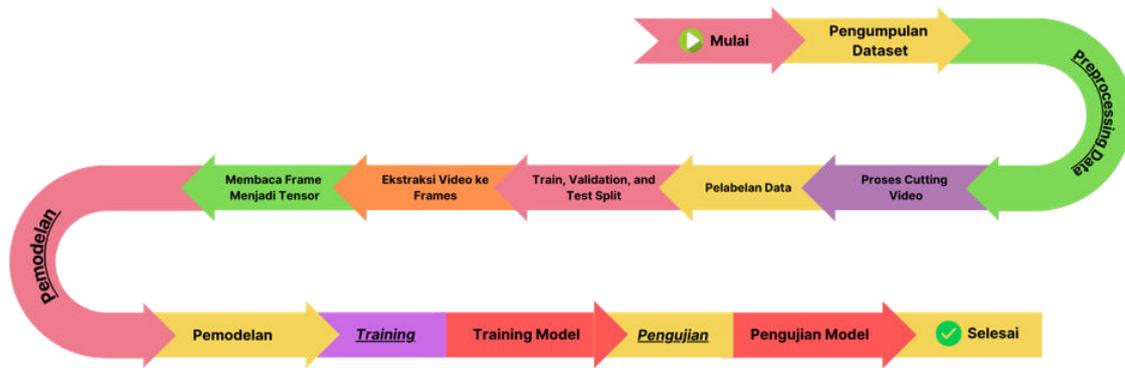


Gambar 1. Penempatan Kamera

2. Pre-processing

Tahap selanjutnya adalah *pre-processing* yang terdiri dari proses *cutting* video. Proses *cutting* video memisahkan satu video yang berisi satu jenis gerakan yang diulang sebanyak lima kali menjadi lima video terpisah. Setelah proses selesai dilakukan, maka jumlah keseluruhan video yang dihasilkan dari satu tempat pengambilan video sebanyak 90 video × 5 pengulangan = 450 video. Lihat Gambar 3 untuk rincian pengambilan video. Berdasarkan penjelasan sebelumnya disampaikan bahwa pengambilan video dilakukan di dua tempat berbeda, maka jumlah dataset video yang didapatkan oleh peneliti sebanyak 900 video.

Kemudian tahap berikutnya adalah membagi data secara random menjadi *train set*, *validation set*, dan *test set*. Proporsi pembagian data dapat dijelaskan pada langkah berikut, pertama, seluruh dataset dibagi dengan proporsi 80% *training set* dan 20% untuk *test set*, dengan jumlah video sebanyak 720 video dan 180 video secara berturut-turut. Kedua, *validation set* kemudian diambil dari *training set* dengan proporsi sebesar 20% juga, maka saat ini jumlah *training set* berubah menjadi 576 video dan jumlah *validation set* sebanyak 144 video.



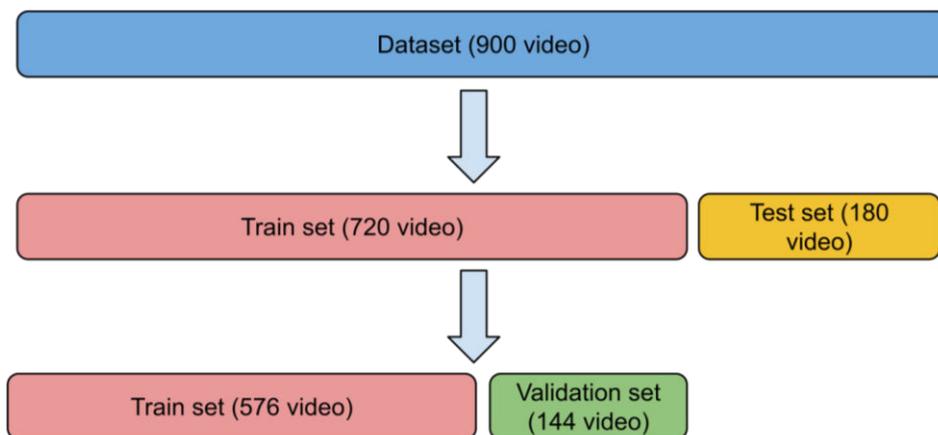
Gambar 2. Alur Penelitian

GETARI

- 90 video
- 6 jenis gerakan tari
- 5 penari (5 perempuan)
- 3 kamera (3 sudut pengambilan)

Gerakan	Penari	Koreografi	Kamera	Pengulangan
Gerakan Ngeed 15 video (5 penari × 3 kamera)	1	1 gerakan per penari	3	5 kali per gerakan
Gerakan Ngelo 15 video (5 penari × 3 kamera)	1	1 gerakan per penari	3	5 kali per gerakan
Gerakan Ngelung 15 video (5 penari × 3 kamera)	1	1 gerakan per penari	3	5 kali per gerakan
Gerakan Ngeseh 15 video (5 penari × 3 kamera)	1	1 gerakan per penari	3	5 kali per gerakan
Gerakan Ngumbang 15 video (5 penari × 3 kamera)	1	1 gerakan per penari	3	5 kali per gerakan
Gerakan Tapak Sirang Pada 15 video (5 penari × 3 kamera)	1	1 gerakan per penari	3	5 kali per gerakan

Gambar 3. Teknik Pengambilan Video



Gambar 4. Pembagian Dataset

Tabel 1. Rincian Data Training Per Kelas

Kelas	Jumlah Data Training	Jumlah Data Validation	Jumlah Data Testing
Ngumbang	96	24	30
Ngeed	96	24	30
Tapak Sirangpada	96	24	30
Ngelo	96	24	30
Ngeseh	96	24	30
Ngelung	96	24	30

Ilustrasi proses pembagian ini dapat dilihat pada Gambar 4. Rincian jumlah data per kelas/label pada *training set* dapat dilihat pada Tabel 1, terlihat pada tabel bahwa proses pembagian ini dapat menghasilkan data yang cukup seimbang per kelasnya.

Sebelum memasukkan *dataset* ke dalam model, hal yang perlu dilakukan adalah memproses data video sedemikian rupa sehingga struktur data video tersebut bersesuaian dengan struktur data yang digunakan pada input model. Sebagai contoh, secara bawaan, struktur data input yang diterima oleh model VGG pada *framework* Tensorflow adalah 224×224 dengan format panjang \times lebar. Jika struktur data *dataset* memiliki *channel* warna Red-Green-Blue (RGB) maka struktur data inputnya menjadi (224, 224, 3), dengan format (panjang *frame*, lebar *frame*, *channel* warna *frame*). Disamping itu, jumlah *frame* yang dimasukkan ke dalam model harus ditetapkan. Secara umum, semakin banyak jumlah *frame* yang per video yang dimasukkan maka semakin banyak pengetahuan yang dapat dipelajari oleh model, namun semakin tinggi sumber daya yang diperlukan saat proses *training*, begitu juga sebaliknya jika menggunakan jumlah *frame* yang sedikit. Untuk menghitung jumlah *frame* yang dimasukkan ke dalam video menggunakan Persamaan 1 dan 2 berikut:

$$frame_{total} = f_{rate} \times V_l \quad (1)$$

$$frame_{step} = \frac{frame_{total}}{frame_n} \quad (2)$$

Keterangan:

- $frame_{total}$: jumlah *frame* dalam satu video
- f_{rate} : *frame rate* video
- V_l : durasi video dalam detik
- $frame_{step}$: langkah *sampling frame*
- $frame_n$: *sample frame*

Asumsikan satu video dengan f_{rate} 30 *frame* per detik, durasi video adalah 5 detik, maka jumlah keseluruhan *frame* dalam satu video sebanyak 150. Maka, $frame_{step}$ yang didapatkan adalah 5. Sehingga, proses pengambilan *frame* akan dilakukan berdasarkan kelipatan nilai $frame_{step}$. Secara keseluruhan parameter *dataset* video yang dimasukkan ke dalam model dapat dilihat pada Tabel 2. Rincian keseluruhan jumlah *frame* yang dihasilkan untuk setiap kelas pada masing-masing data ditunjukkan pada Tabel 3.

Tabel 2. Parameter Data Video

Parameter	Nilai
-----------	-------

Ukuran batch	2
Jumlah channel warna	3
Jumlah frame per video	30
Panjang frame	224
Lebar frame	224

Tabel 3. Detail Jumlah Frame Per Kelas

Kelas	Total Jumlah Frame		
	Data Training	Data Validation	Data Test
Ngumbang	2880	720	900
Ngeed	2880	720	900
Tapak	2880	720	900
Sirangpada	2880	720	900
Ngelo	2880	720	900
Ngeseh	2880	720	900
Ngelung	2880	720	900

3. Pemodelan

Pada bagian ini dijelaskan model yang digunakan sebagai klasifier. Peneliti menggunakan model *pre-trained* VGG16 (yaitu VGG dengan jumlah *layer* konvolusi sebanyak 16 *layer*) [20] dan Inflated 3D Network (I3D) [21]. Kedua model *pre-trained* ini digunakan dalam proses *feature engineering*. VGG16 dan I3D diimplementasikan dengan *framework* deep learning Tensorflow, di mana VGG16 menggunakan *weights* dari *dataset* ImageNet [23], sedangkan I3D menggunakan *weights* dari *dataset* Kinetics-400 [24].

Perbedaan mendasar dari model *pre-trained* VGG dan I3D adalah dari pendekatan dalam memproses *frame* di dalam video. VGG pada dasarnya dirancang sebagai klasifier data citra saja, maka VGG memproses individual citra atau *frame* ke dalam layernya kemudian menghitung rata-rata probabilitas akhir kelas dari vektor citranya. Berbeda dengan VGG, I3D menggunakan konvolusi tiga dimensi sebagai teknik belajar informasi *spatio-temporal* secara langsung dalam suatu data video [25]. I3D dikembangkan untuk meningkatkan kemampuan Convolutional 3D Networks (C3D) [26] di mana pada dasarnya dirancang sebagai klasifier video terutama pada permasalahan *action recognition*. C3D dilatih menggunakan *dataset* Sports-1M [25].

Seperti yang telah dijelaskan pada bagian sebelumnya, model *pre-trained* yang digunakan berfungsi sebagai *feature engineer* saja maka perlu ditambahkan *custom layer* yang berfungsi sebagai klasifiernya. *Custom layer* ini berdasarkan dari penelitian sebelumnya [17]. Berikut Gambar 5 dan Gambar 6 adalah arsitektur keseluruhan model yang digunakan.

Layer paling atas pada arsitektur VGG16-LSTM (Gambar 5) adalah *layer input* yang merupakan video berupa *frames* sebanyak

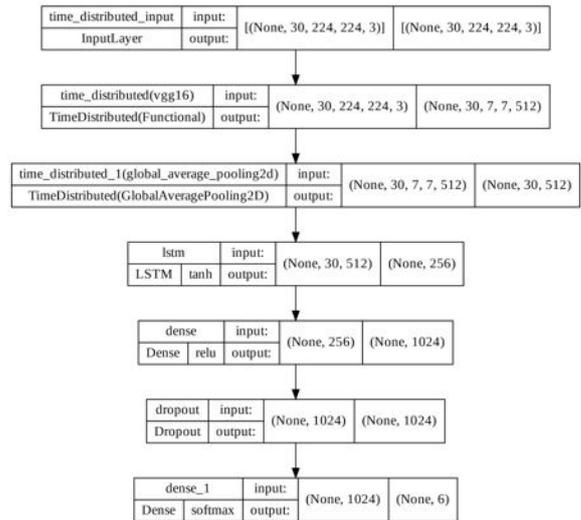
30 dengan ukuran gambar 224×224 dan 3 *channel* warna (RGB). Setelah *layer input* dilanjutkan dengan arsitektur model VGG16 [20] itu sendiri. Model VGG16 yang digunakan yaitu *pre-trained* model yang disediakan oleh framework TensorFlow Keras. *Layer* VGG16 berada di dalam *layer* TimeDistributed. Setelah *layer* VGG16 terdapat *layer* GlobalAveragePooling2D yang juga berada di dalam *layer* TimeDistributed. *Layer* selanjutnya adalah *layer* LSTM dengan jumlah *neuron* sebanyak 256 *units*. Setelah *layer* LSTM terdapat *Dense layer* sebanyak 1024 *neuron* dengan fungsi aktivasi Rectified Linear Unit (ReLU). Setelah *Dense layer* terdapat *layer* Dropout yang berfungsi untuk mengurangi resiko overfit. Setelah *layer* Dropout, terdapat *layer* Dense sebanyak 7 *neuron* yang mewakili tujuh gerakan dasar Tari Bali dengan fungsi aktivasi Softmax untuk menormalisasi hasil akhir dari klasifikasi.

Serupa dengan model VGG16-LSTM, *Layer* paling atas dari I3D adalah *layer input* yang merupakan video berupa *frames* sebanyak 30 dengan ukuran gambar 224×224 dan 3 *channel* warna (RGB). Setelah *layer input*, *layer* selanjutnya adalah *pre-trained* model I3D [21]. Model I3D pun menggunakan *pre-trained* model yang telah diimplementasikan ke framework TensorFlow Keras. Setelah *layer* I3D terdapat *layer* GlobalAveragePooling3D. Kemudian dilanjutkan dengan *layer* Dense sebanyak 1024 *neuron* dengan fungsi aktivasi ReLU, setelah itu *layer* Dense dan *Dropout layer*. Setelah *Dropout layer*, terdapat *Dense layer* sebanyak 7 *neuron* output.

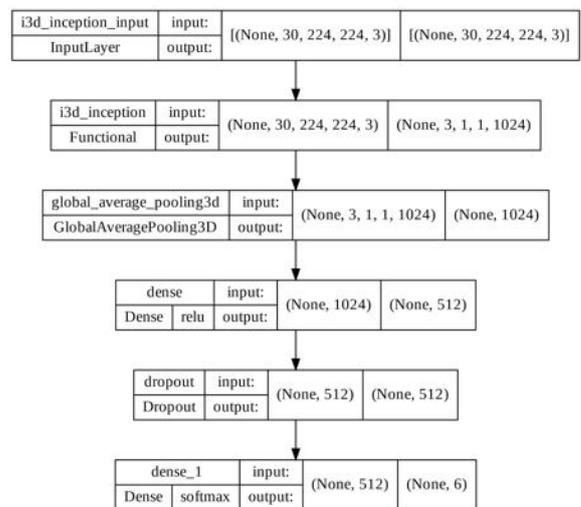
HASIL DAN PEMBAHASAN

A. Hasil

Hal yang perlu ditentukan sebelum memulai proses training adalah menentukan nilai-nilai *hyper-parameter* seperti fungsi *loss*, fungsi optimasi, jumlah *epoch*, teknik *early stopping*, dan metrik pengujian.



Gambar 5. Arsitektur VGG16-LSTM



Gambar 6. Arsitektur I3D

Saat proses training, model mempelajari data berdasarkan jumlah *epoch* yang sudah ditentukan, yaitu sebanyak 25 *epoch*. Dengan teknik *early stopping*, model tidak perlu menempuh seluruh *epoch* untuk mencapai kondisi konvergen. Jika dalam jumlah *epoch* tertentu secara berturut-turut tidak ada penurunan nilai *validation loss* maka proses training akan langsung dihentikan. Penentuan jumlah *epoch* pada konteks ini dapat ditentukan dengan menentukan parameter *patience*, di mana pada penelitian ini ditentukan sebesar 3. Teknik *early stopping* sangat umum digunakan sebagai strategi *training* untuk menghemat penggunaan sumber daya komputasi.

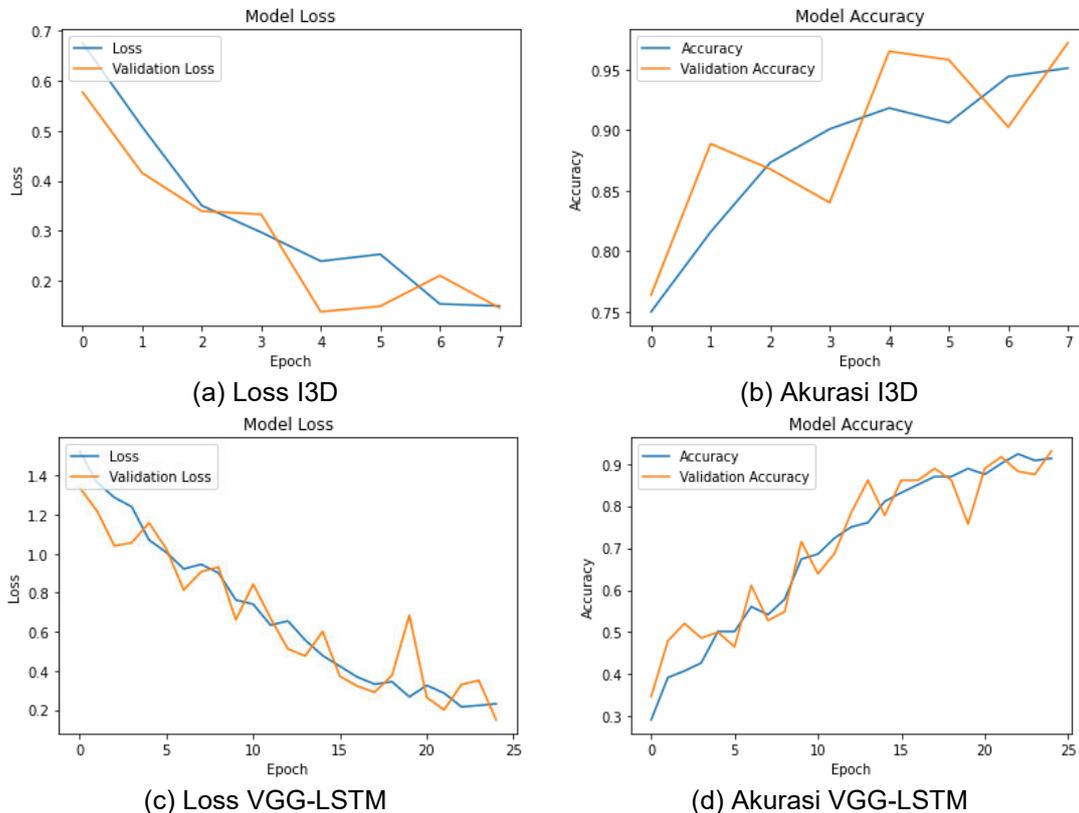
Berdasarkan hasil training pada model VGG16-LSTM didapatkan nilai *training loss* sebesar 0,23 dan akurasi sebesar 0,91 (Gambar 7.c dan 7.d). Pada *data validation* didapatkan nilai *validation loss* sebesar 0,14 dan akurasi sebesar 0,93. Nilai ini didapatkan pada proses

training yang mencapai 25 *epoch*. Walaupun menggunakan teknik *early stopping*, model VGG16-LSTM tetap mengalami perubahan nilai *validation loss* hingga *epoch* ke-25. Hasil yang didapatkan oleh model VGG16-LSTM sedikit tertinggal dibandingkan dengan yang dihasilkan oleh I3D. I3D memperoleh nilai *training loss* sebesar 0,14 dengan nilai akurasi mencapai 0,95. Begitu juga dengan nilai yang didapatkan dari pengujian data validasi, di mana *validation loss* yang dihasilkan sebesar 0,14 dan juga nilai akurasi sebesar 0,97. Proses *training* model I3D berhenti pada *epoch* ke-7. Grafik *loss* dan akurasi dari model I3D dapat dilihat pada Gambar 7.a dan 7.b secara berturut-turut.

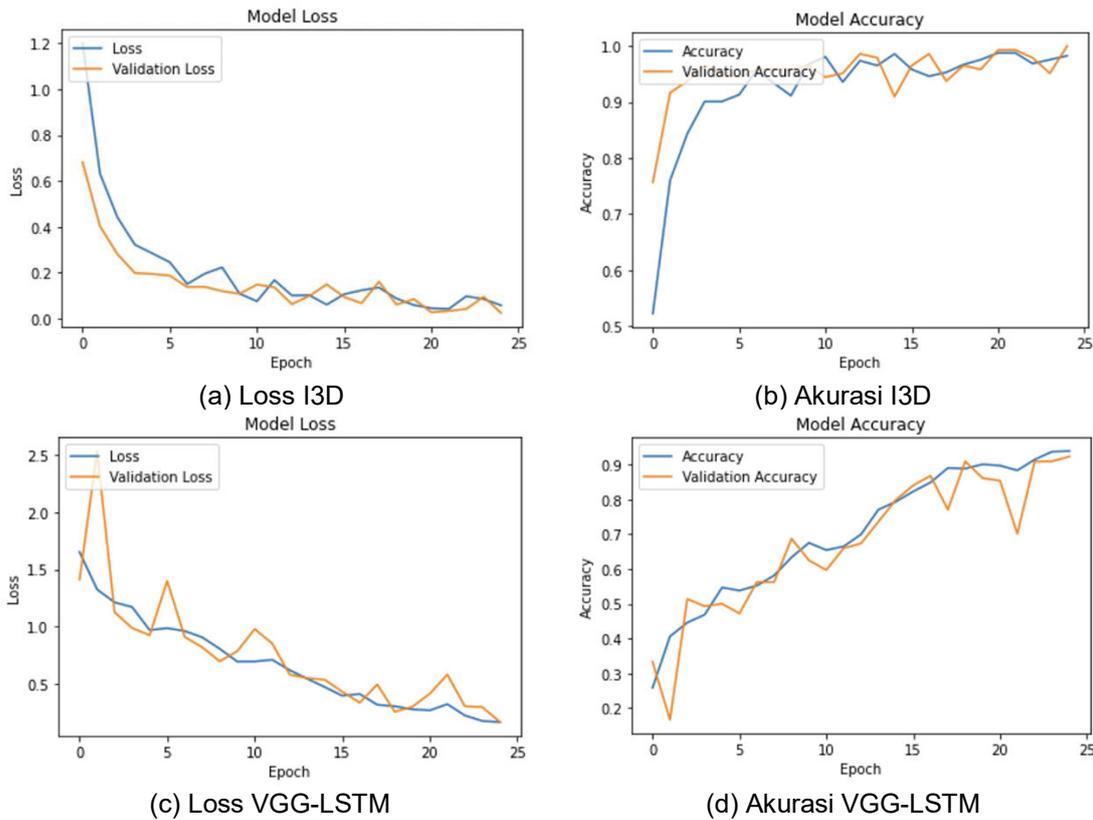
dan model VGG16-LSTM menghasilkan *validation loss* dengan nilai yang sama sebesar 0,14. Sedangkan proses *training* pada VGG16-LSTM berhenti pada saat *epoch* terakhir (lihat Gambar 7.c dan d). Berdasarkan hasil ini dapat dikatakan bahwa VGG16-LSTM memerlukan *epoch* yang lebih banyak untuk mencapai keadaan konvergen daripada model I3D. Oleh karena itu, pada penelitian ini dilakukan eksperimen yang kedua. Pada eksperimen kedua (lihat Gambar 8), kedua model di-training dengan jumlah *epoch* sebanyak 25 dan tanpa menggunakan teknik *early stopping*. Hal ini dilakukan dengan tujuan untuk mengetahui bagaimana kinerja model jika diberikan jumlah *epoch* yang sama.

B. Pembahasan

Berdasarkan pemaparan sebelumnya terlihat jelas bahwa kinerja model VGG16-LSTM berada di bawah model I3D. Model I3D menghasilkan nilai *validation loss* sebesar 0,14



Gambar 7. Loss dan Akurasi Eksperimen 1



Gambar 8. Loss dan Akurasi Eksperimen 2

Pada eksperimen kedua, nilai loss model I3D menurun tajam hingga mencapai 0,05, bahkan nilai validation loss mencapai 0,02. Sedangkan pada model VGG-LSTM tidak ada perubahan yang signifikan pada *loss* dan juga *validation loss*.

Pengambilan nilai metrik pengukuran lainnya seperti *precision*, *recall*, dan F1 juga dilakukan pada penelitian ini. Alasan peneliti melakukan pengukuran juga menggunakan keempat parameter ini adalah sebagai berikut:

- 1) Dataset yang digunakan terdiri dari enam kelas, maka permasalahan yang dipecahkan adalah klasifikasi multi kelas.
- 2) Pentingnya mengetahui keseimbangan kinerja model dalam pengklasifikasian per kelas data. Keseimbangan kinerja model dalam hal ini terkait dengan tingkat bias model terhadap hasil prediksi.
- 3) Tingkat bias dapat dideteksi langsung dari nilai klasifikasi *precision*, *recall*, ataupun F1.

Seperti yang telah dijelaskan sebelumnya, ada penambahan eksperimen untuk mengetahui kinerja model. Pada Gambar 8, terlihat perubahan yang cukup signifikan pada nilai *loss* dan akurasi model I3D, sedangkan *loss* dan akurasi pada model VGG-LSTM tidak terlihat seperti ada perubahan yang berarti. Tabel 4

menunjukkan kinerja model pada *data test* baik VGG16-LSTM dan I3D pada eksperimen pertama dan kedua dengan 25 *epoch* pada seluruh nilai metrik. Ketika membandingkan kinerja model I3D dan VGG-LSTM pada eksperimen kedua barulah terlihat perbedaan yang cukup signifikan, hal ini diperkuat pada grafik yang ditunjukkan pada Gambar 9.

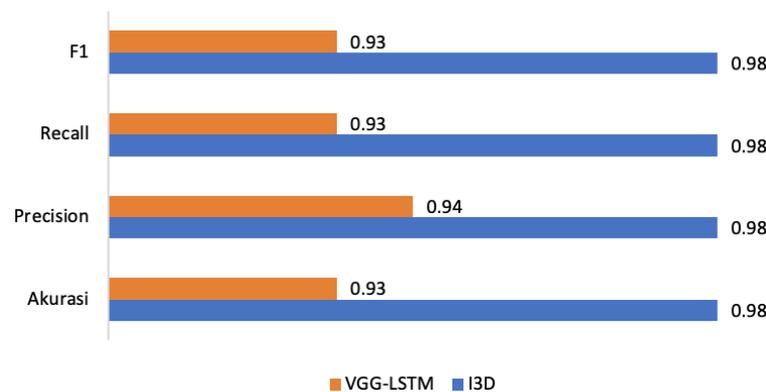
Jika ditinjau berdasarkan permasalahan klasifikasi multi kelas seperti yang disampaikan sebelumnya, pada alasan peneliti poin 1 nilai *precision*, *recall*, dan F1 pada Tabel 4 menunjukkan bahwa kinerja kedua model sangat baik. Hal ini menunjukkan bahwa dataset yang dibangun memiliki proporsi jumlah kelas yang seimbang.

Nilai yang ditunjukkan pada Tabel 4, diperkuat oleh hasil pengukuran metrik yang ditunjukkan pada Tabel 5. Pada tabel tersebut, disajikan nilai metrik *precision*, *recall*, dan F1 pada kelas data *ngelung*, *ngeseh*, *tapak sirangpada*, *ngeed*, *ngelo*, dan *ngumbang*. Nilai metrik pada kedua model berkisar antara 0,82 hingga 1. Nilai-nilai ini menunjukkan bahwa kedua model memiliki kemampuan yang baik dalam membedakan atau mengklasifikasikan masing-masing kelas data. Sehingga dapat dikatakan bahwa hasil klasifikasi dari model tidak bias.

Tabel 4. Nilai Metrik pada Data Test

Metrik	Eksperimen 1		Eksperimen 2	
	I3D	VGG-LSTM	I3D	VGG-LSTM
Akurasi	0,95	0,93	0,98	0,93
Precision	0,96	0,94	0,98	0,94
Recall	0,95	0,93	0,98	0,93
F1	0,95	0,93	0,98	0,93

I3D vs VGG-LSTM pada Data Test



Gambar 9. Grafik Kinerja Model pada Data Test

Tabel 5. Nilai Metrik Per Kelas pada Data Test

Kelas	Precision		Recall		F1	
	VGG-LSTM	I3D	VGG-LSTM	I3D	VGG-LSTM	I3D
Ngeed	0,87	0,97	0,96	1	0,92	0,98
Ngelo	1	1	1	0,97	1	0,99
Ngelung	0,94	0,97	0,86	0,94	0,9	0,96
Ngeseh	0,94	0,95	0,89	0,97	0,92	0,96
Ngumbang	0,82	1	0,9	1	0,86	1
Tapak Sirangpada	1	1	1	1	1	1

KESIMPULAN

Berdasarkan pemaparan yang telah disampaikan pada bagian sebelumnya maka kesimpulan pada penelitian ini adalah telah dibuat dataset gerakan dasar Tari Bali yang dinamakan GETARI, meliputi gerakan *ngelung*, *ngeseh*, *tapak sirangpada*, *ngeed*, *ngelo*, dan *ngumbang*.

Hasil klasifikasi menggunakan dataset GETARI dengan model *pre-trained* VGG16-LSTM dan I3D menunjukkan bahwa dataset ini memiliki keseimbangan data yang baik. Hal ini ditunjukkan berdasarkan hasil pengukuran secara keseluruhan pada metrik *precision*, *recall*, dan F1 pada data test. Nilai *precision*, *recall*, dan F1 pada model VGG-LSTM yang cukup merata

baik pada eksperimen pertama dan kedua. Begitu juga nilai metrik pengukuran dari model I3D sangat merata disemua metrik pengujian. Hasil ini diperkuat dengan nilai metrik yang disajikan berdasarkan masing-masing kelas pada dataset. Nilai-nilai ini menunjukkan bahwa kedua model sama sekali tidak mengalami kesulitan dalam mengklasifikasikan masing-masing kelas data.

MASUKAN PENELITIAN

Pada penelitian ini dataset yang dibuat belum dapat merepresentasikan problem set yang lengkap untuk seluruh gerakan dasar Tari Bali perempuan (13 gerakan) dan juga gerakan dasar Tari Bali laki-laki. Hal ini dikarenakan

sampai pada artikel ini ditulis peneliti belum memiliki sumber daya komputasi yang mencukupi. Oleh karena itu, pada penelitian selanjutnya peneliti akan menyertakan seluruh gerakan tari untuk penari perempuan dan juga membuat dataset gerakan tari untuk laki-laki.

UCAPAN TERIMAKASIH

Penelitian ini didanai oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, dan Institut Teknologi dan Bisnis STIKOM Bali melalui Program Kreativitas Mahasiswa (PKM) Riset Eksakta skema pendanaan tahun 2022.

REFERENSI

- [1] I. L. Koten and C. R. A. Pramatha, "Semantic Representation of Balinese Traditional Dance," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 8, no. 4, p. 411, Feb. 2020, doi: 10.24843/JLK.2020.V08.I04.P07.
- [2] A. W. R. Emanuel and A. Widjaja, "Feasibility study of scripting Indonesian traditional dance motion in XML format," *Proceedings - 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2017*, vol. 2018-January, pp. 57–61, Feb. 2018, doi: 10.1109/ICITISEE.2017.8285559.
- [3] N. U. Januhari, N. L. A. K. Y. Sarja, and M. Rudita, "Information Technology Based Model for Balinese Classical Dance Preservation," *International Journal of Engineering Technologies and Management Research*, vol. 5, no. 11, pp. 71–85, 2018, doi: 10.5281/zenodo.2280898.
- [4] N. Anbarsanti and A. S. Prihatmanto, "Dance modelling, learning and recognition system of aceh traditional dance based on hidden Markov model," *2014 International Conference on Information Technology Systems and Innovation, ICITSI 2014 - Proceedings*, pp. 86–92, Feb. 2014, doi: 10.1109/ICITSI.2014.7048243.
- [5] N. Anbarsanti and A. S. Prihatmanto, "HMM-based model for dance motions with pose representation," May 2014. doi: 10.1109/ICSENGT.2014.7111793.
- [6] N. Anbarsanti and A. S. Prihatmanto, "Dance learning and recognition system based on hidden Markov model. a case study: Aceh traditional dance," May 2014. doi: 10.1109/ICSENGT.2014.7111792.
- [7] H. Hoettinger, F. Mally, and A. Sabo, "Activity Recognition in Surfing - A Comparative Study between Hidden Markov Model and Support Vector Machine," *Procedia Eng*, vol. 147, pp. 912–917, Jan. 2016, doi: 10.1016/J.PROENG.2016.06.279.
- [8] P. Asghari, E. Soleimani, and E. Nazerfard, "Online Human Activity Recognition Employing Hierarchical Hidden Markov Models," *J Ambient Intell Humaniz Comput*, vol. 11, no. 3, pp. 1141–1152, Mar. 2019, doi: 10.48550/arxiv.1903.04820.
- [9] M. Panzner and P. Cimiano, "Comparing hidden Markov models and long short term memory neural networks for learning action representations," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10122 LNCS, pp. 94–105, 2016, doi: 10.1007/978-3-319-51469-7_8/COVER.
- [10] Y. Heryadi, M. I. Fanany, and A. Murni, "A skeleton descriptor for kinesthetic element recognition of bali traditional dances," *Proceedings of International Conference on Computational Intelligence, Modelling and Simulation*, pp. 43–47, 2012, doi: 10.1109/CIMSIM.2012.78.
- [11] Y. Heryadi, M. I. Fanany, and A. M. Arymurthy, "A syntactical modeling and classification for performance evaluation of Bali traditional dance," in *2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2012, pp. 261–265.
- [12] Y. Heryadi, M. I. Fanany, and A. M. Arymurthy, "Grammar of Dance Gesture from Bali Traditional Dance," *Int J Comp Sci*, vol. 9, no. 6, pp. 144–149, Nov. 2012, Accessed: Feb. 22, 2022. [Online]. Available: www.IJCSI.org
- [13] M. W. A. Kesiman, I. M. D. Maysanjaya, I. M. A. Pradnyana, I. M. G. Sunarya, and P. H. Suputra, "Revealing the Characteristics of Balinese Dance Maestros by Analyzing Silhouette Sequence Patterns Using Bag of Visual Movement with HoG and SIFT Features," *Journal of ICT Research and Applications*, vol. 15, no. 1, pp. 89–104, Jul. 2021, doi: 10.5614/ITBJ.ICT.RES.APPL.2021.15.1.6.
- [14] L. Zaman, S. Sumpeno, M. Hariadi, Y. Kristian, E. Setyati, and K. Kondo,

- “Modeling Basic Movements of Indonesian Traditional Dance Using Generative Long Short-Term Memory Network,” *IAENG Int J Comput Sci*, vol. 47, no. 2, Jun. 2020, Accessed: Mar. 09, 2022. [Online]. Available: http://www.iaeng.org/IJCS/issues_v47/issue_2/IJCS_47_2_14.pdf
- [15] L. Zaman, S. Sumpeno, and M. Hariadi, “Training Strategies for Remo Dance on Long Short-Term Memory Generative Model,” in *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia*, Jul. 2018, pp. 176–180. doi: 10.1109/CENIM.2018.8710992.
- [16] L. Zaman, S. Sumpeno, and M. Hariadi, “Analisis Kinerja LSTM dan GRU sebagai Model Generatif untuk Tari Remo,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, vol. 8, no. 2, pp. 142–150, May 2019, doi: 10.22146/JNTETI.V8I2.503.
- [17] P. A. T. Permana and I. N. R. Hendrawan, “Klasifikasi Video Gerakan Dasar Tari Bali Berbasis Machine Learning Dengan Neural Network,” Denpasar, Dec. 2021.
- [18] K. D. B. B. P. Bali, *Pendokumentasian Gerak-Gerak Tari Bali*, 1st ed. Dinas Perpustakaan dan Kearsipan Kota Denpasar, 2000. [Online]. Available: <https://perpustakaan.denpasarkota.go.id/opac/detail-opac?id=16282>
- [19] Team Survey ASTI Denpasar, *Gerak Tari Bali*. Denpasar: Akademi Seni Tari Indonesia Denpasar, 1983.
- [20] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Sep. 2014, Accessed: Feb. 22, 2022. [Online]. Available: <https://arxiv.org/abs/1409.1556v6>
- [21] J. Carreira, A. Zisserman, Z. Com, and † Deepmind, “Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 4724–4733, May 2017, doi: 10.1109/CVPR.2017.502.
- [22] S. Tsuchida, S. Fukayama, M. Hamasaki, and M. Goto, “AIST Dance Video Database: Multi-Genre, Multi-Dancer, and Multi-Camera Database for Dance Information Processing,” 2019.
- [23] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” pp. 248–255, Mar. 2010, doi: 10.1109/CVPR.2009.5206848.
- [24] W. Kay et al., “The Kinetics Human Action Video Dataset,” May 2017, doi: 10.48550/arxiv.1705.06950.
- [25] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, “Large-Scale Video Classification with Convolutional Neural Networks,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 1725–1732. doi: 10.1109/CVPR.2014.223.
- [26] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning Spatiotemporal Features with 3D Convolutional Networks,” in *Proceedings of the IEEE International Conference on Computer Vision*, Dec. 2014, pp. 4489–4497. Accessed: Sep. 24, 2022. [Online]. Available: <https://arxiv.org/abs/1412.0767v4>