

Perbandingan Algoritma Convolutional Neural Network dan Capsule Network Dalam Klasifikasi Jenis Rumah Adat

Yosefina Finsensia Riti^{1*}, Yulia Wahyuningsih², Josephine Roosandriantini³, Paulus William Siswanto⁴

^{1,2,4} Program Studi Ilmu Informatika, Universitas Katolik Darma Cendika, Surabaya, Jawa Timur

³ Program Studi Arsitektur, Universitas Katolik Darma Cendika, Surabaya, Jawa Timur

Email: ^{1*} yosefina.riti@ukdc.ac.id, ² yulia@ukdc.ac.id, ³ jose.roo@ukdc.ac.id, ⁴ paulus.siswanto@student.ukdc.ac.id

(Naskah masuk: 3 Okt 2023, direvisi: 31 Okt 2023, diterima: 6 Nov 2023)

Abstrak

Warisan budaya merupakan salah satu bagian yang penting dalam budaya Indonesia dan perlu dilestarikan keberadaannya, salah satu warisan budaya yang perlu dilestarikan yaitu Rumah adat. Informasi terkait rumah adat juga cenderung dicari oleh masyarakat untuk kepentingan pendidikan dalam dunia pengenalan budaya, arsitektur, maupun dalam bidang pariwisata. Hingga saat ini masih banyak daerah di Indonesia yang masih memiliki rumah adatnya dengan tujuan untuk mempertahankan nilai budaya, sebagai tempat pertemuan acara adat, maupun sebagai objek wisata. Salah satunya adalah Provinsi Nusa Tenggara Timur (NTT) yang memiliki beragam rumah adat seperti Sao Ngada (Suku Bajawa), Sao Ria (Suku Ende), Ume Kbbubu (Suku Timor), Sao Ata Mosa Lakitana (Sumba), dan Mbaru Niang (Suku Wae Rebo). Keanekaragaman bentuk, kemiripan material penyusun, dan kemiripan bangunan dapat membuat sebagian masyarakat awam kesulitan dalam membedakan jenis atau nama antara rumah adat yang satu dengan rumah adat yang lain. Oleh karena itu diperlukan teknologi digital yang dapat mengidentifikasi dan mengklasifikasikan rumah adat, sehingga dapat membantu wisatawan, maupun masyarakat umum yang mempelajari seputar rumah adat dalam membedakan jenis rumah adat tertentu. Dalam Penelitian ini model *deep learning* diterapkan untuk identifikasi dan klasifikasi rumah adat, dengan menggunakan algoritma *Convolutional Neural Network* (CNN) dan *Capsule Networks* (*CapsNet*). Penelitian ini bertujuan membandingkan algoritma *deep learning*, CNN menggunakan arsitektur *Resnet50V2* dan *CapsNet*, dimana *dataset* yang digunakan terdiri dari 500 data rumah adat di NTT. Hasil penelitian menunjukkan bahwa CNN memiliki nilai akurasi sekitar 98% dengan nilai *loss* sekitar 0,72, sedangkan *CapsNet* memiliki nilai akurasi sekitar 72% dengan nilai *loss* sekitar 1,73%. Berdasarkan hasil klasifikasi tersebut dalam kasus ini, disimpulkan bahwa CNN lebih baik dalam mengklasifikasikan objek rumah adat dibandingkan dengan *CapsNet*. Oleh karena itu untuk pekerjaan lebih lanjut dapat dilakukan perbaikan parameter *tuning* untuk algoritma *CapsNet* dan juga dapat mengimplementasikan CNN dalam pembuatan aplikasi untuk identifikasi dan klasifikasi objek rumah adat sehingga dapat membantu Masyarakat umum dalam membedakan jenis rumah adat melalui aplikasi tersebut.

Kata Kunci: *Deep Learning*, CNN, *CapsNet*, Klasifikasi, Warisan Budaya.

Abstract

Cultural heritage is one of the important aspects of Indonesian culture that needs to be preserved. One of the cultural heritages that need to be preserved is traditional houses. Information related to traditional houses is also often sought by the community for educational purposes in the fields of culture, architecture, and tourism. To this day, many regions in Indonesia still have their traditional houses with the purpose of maintaining cultural values, as venues for traditional events, and as tourist attractions. One of these regions is East Nusa Tenggara Province (NTT), which has a variety of traditional houses such as Sao Ngada (Bajawa Tribe), Sao Ria (Ende Tribe), Ume Kbbubu (Timor Tribe), Sao Ata Mosa Lakitana (Sumba), and Mbaru Niang (Wae Rebo Tribe). The diversity of forms, similarity of constituent materials, and building similarities can make it difficult for the general public to distinguish between different types or names of traditional houses. Therefore, digital technology is needed to identify and classify traditional houses, helping tourists and the general public learn about different types of traditional houses. In this research, a deep learning model is applied for the identification and classification of traditional houses, using Convolutional Neural Network (CNN) and Capsule Networks (CapsNet) algorithms. This research aims to compare the deep learning algorithms, specifically CNN using the Resnet50V2 architecture and CapsNet. The dataset used consists of 500 traditional house images from NTT. The research results show that CNN has an accuracy of around 98% with a loss value of approximately 0.72, while CapsNet has an accuracy of about 72% with a loss value of around 1.73. Based on these classification

results, it is concluded that CNN is better at classifying traditional houses compared to CapsNet. Therefore, for further work, parameter tuning for the CapsNet algorithm can be improved, and CNN can be implemented in the development of an application for the identification and classification of traditional houses to assist the general public in distinguishing between different types of traditional houses through the application.

Keywords: Deep Learning, CNN, CapsNet, Classification, Culture Heritage.

I. PENDAHULUAN

Salah satu aspek yang sangat penting dalam menjaga warisan budaya Indonesia adalah rumah adat. Rumah adat adalah bangunan yang memiliki karakteristik unik dan dihuni oleh suku-suku tertentu, dengan perbedaan dalam bentuk atap, bahan bangunan, arsitektur, dan tata ruang internal [1]. Hingga hari ini, masih banyak daerah di Indonesia yang mempertahankan rumah adat mereka dengan tujuan untuk melestarikan nilai-nilai budaya, sebagai tempat untuk acara adat, dan juga sebagai objek wisata [2]. Nusa Tenggara Timur (NTT) adalah salah satu provinsi di Indonesia yang memiliki kekayaan rumah adat, termasuk rumah adat Sao Ria (Suku Ende), Mbaru Niang (Suku Wae Rebo), Sao Ata Mosa Lakitana (Sumba), Sao Ngada (Suku Bajawa), dan Ume Kbbu (Suku Timor). Informasi tentang rumah adat sering dicari oleh masyarakat untuk kepentingan pendidikan tentang budaya, arsitektur, dan juga dalam konteks pariwisata.

Namun, ada beberapa permasalahan yang dihadapi dalam mengenali rumah adat, seperti variasi bentuk yang beragam, kesamaan bahan bangunan, dan kemiripan struktur bangunan, yang dapat menyulitkan dalam membedakan jenis rumah adat satu dengan yang lainnya. Oleh karena itu, diperlukan teknologi digital yang dapat mengidentifikasi dan mengklasifikasikan rumah adat, sehingga dapat membantu wisatawan dan masyarakat umum dalam mempelajari dan membedakan jenis rumah adat tertentu. Klasifikasi rumah adat dilakukan dengan tujuan untuk meningkatkan manajemen data gambar dan pencarian spesifik terkait rumah adat di NTT. Teknologi *deep learning* menjadi salah satu metode yang digunakan dalam melakukan klasifikasi citra. *Deep learning* merupakan cabang ilmu dari kecerdasan buatan, khususnya dalam bidang *Machine learning*, yang menggunakan konsep jaringan saraf tiruan dengan banyak lapisan tersembunyi. Pendekatan *deep learning* telah diimplementasikan dalam berbagai bidang, termasuk pengenalan objek [3], [4], pengolahan citra [5]–[7], pengenalan suara [8]–[10], citra medis [11]–[13], dan masih banyak lagi.

Terdapat berbagai jenis algoritma *deep learning* yang telah diimplementasikan untuk tujuan klasifikasi objek, diantaranya adalah CNN, *Capsule Networks*, YOLO, *Recurrent Neural Networks* (RNN), dan *Deep Neural Networks* (DNN).

CNN merupakan jenis jaringan saraf tiruan yang dikhususkan untuk memproses citra menggunakan operasi konvolusi, sehingga mampu mengekstrak fitur-fitur penting dari data masukan. CNN banyak digunakan dalam aplikasi

seperti deteksi penyakit medis [12], [14] dan klasifikasi objek [15]. *Capsule Network* merupakan inovasi terbaru yang memperkenalkan "unit pemrosesan" berupa "kapsul" untuk memahami relasi antar fitur dalam citra.

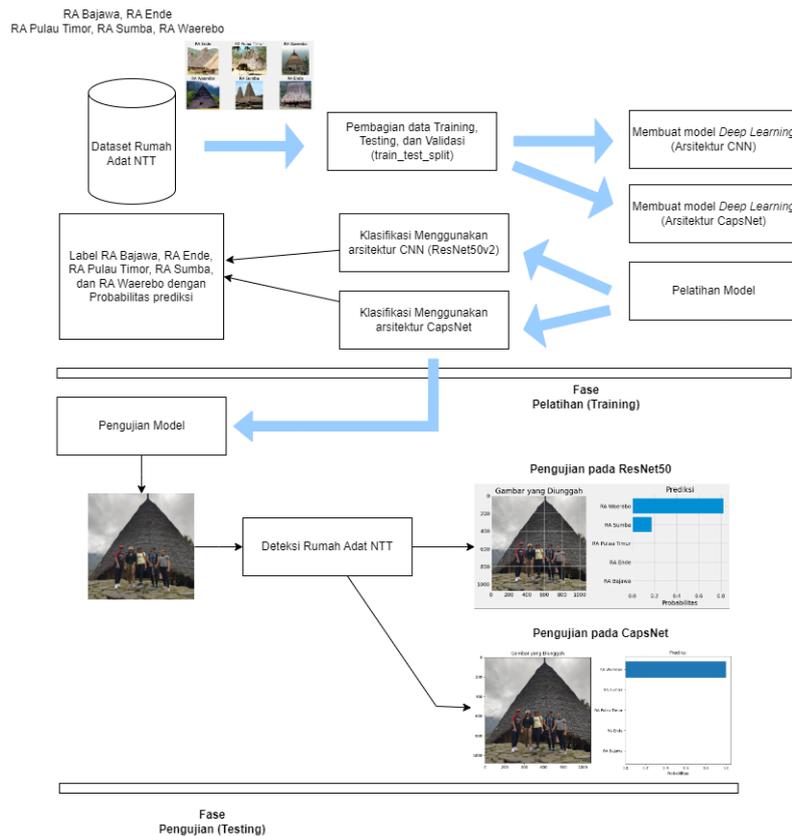
Beberapa penelitian menggunakan CNN untuk mendeteksi *COVID-19*, klasifikasi citra histopatologi, dan kanker kulit [14], [16]–[19] dan terbukti CNN efektif. Di sisi lain beberapa penelitian juga menguji *CapsNet* diantaranya, mengklasifikasikan citra hiperspektral, klasifikasi citra CT paru – paru, dan klasifikasi suara pernapasan [20]–[22].

Sementara untuk klasifikasi citra rumah budaya beberapa peneliti telah mengimplementasikan algoritma CNN, seperti studi Kulkarni dkk. [23] mengklasifikasi situs-situs warisan budaya. Penelitian Cosovic [24] juga menjelajahi penggunaan algoritma jaringan saraf konvolusional dalam mengklasifikasi citra warisan budaya. Sementara itu algoritma *CapsNet* masih belum ada untuk klasifikasi rumah adat. Kinerja Algoritma CNN dan *CapsNet* sudah pernah dibandingkan untuk berbagai jenis klasifikasi citra yaitu dengan menggunakan *dataset MNIST* dan *CIFAR-10*. Dan yang menjadi kebaruan dari penelitian ini dengan penelitian-penelitian sebelumnya adalah melakukan klasifikasi rumah adat dengan membandingkan algoritma CNN dan *CapsNet* dengan *dataset* rumah adat NTT.

Oleh karena itu, dalam penelitian ini merekomendasikan algoritma CNN dan *CapNet* untuk mengklasifikasikan jenis rumah adat daerah NTT, CNN direkomendasikan karena kemampuannya yang kuat dalam mengekstraksi fitur hierarkis dari gambar, terbukti dalam pengenalan objek berdasarkan studi literatur terdahulu. Sedangkan *CapsNet* dipilih untuk mengatasi keterbatasan representasi fitur pada CNN. Kedua algoritma tersebut dipilih karena memiliki performa yang baik dalam mengenali objek maupun mengklasifikasi objek. Tujuan dari penelitian ini adalah klasifikasi citra rumah adat menggunakan algoritma CNN dan *CapsNet* untuk membantu pencarian gambar rumah adat yang sesuai bagi masyarakat umum, pelajar, maupun wisatawan, selain itu dari penelitian ini diperoleh metode klasifikasi dengan kinerja yang paling baik antara CNN dan *CapsNet* dalam melakukan klasifikasi citra rumah adat.

II. METODE PENELITIAN

Metode dalam penelitian ini adalah melakukan pengujian algoritma CNN dan *CapsNet*. Gambar 1 menggambarkan langkah-langkah penelitian ini.

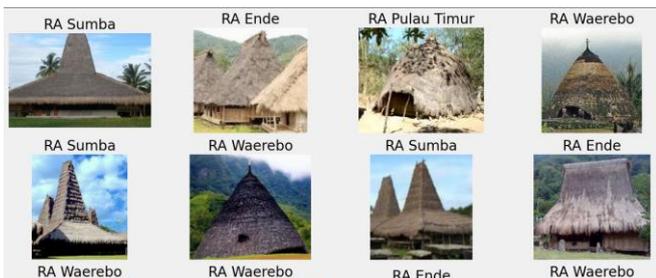


Gambar 1. Flowchart Model yang Diusulkan

A. Pengumpulan Dataset

Pengumpulan *dataset* citra rumah adat diperoleh melalui pengambilan data secara langsung menggunakan kamera dan juga diperoleh dari beberapa sumber internet seperti seperti *Google*, *Bing*, dan media sosial. Pengambilan *dataset* dari internet dilakukan untuk tujuan memperoleh variasi *dataset* rumah adat. Karena rumah adat dari berbagai daerah tersebut diambil pada musim kemarau dan jumlahnya terbatas.

Dataset yang digunakan dalam penelitian ini terdiri dari 500 data rumah adat yang mewakili berbagai jenis rumah adat tradisional dari berbagai suku dan daerah di Indonesia. Setiap kelas dalam *dataset* ini mencakup 100 data yang mewakili jenis rumah adat yang berbeda, yaitu Sao Ngada (Suku Bajawa), Sao Ria (Suku Ende), Ume Kbu (Suku Pulau Timur), Sao Ata Mosa Lakitana (Sumba), dan Mbaru Niang (Suku Wae Rebo).



Gambar 2. Visualisasi Dataset

Gambar 2 adalah contoh visualisasi dari *dataset* rumah adat yang digunakan. Visualisasi ini akan membantu peneliti dalam memahami karakteristik dan perbedaan antara berbagai jenis rumah adat yang terdapat dalam *dataset*, serta berperan sebagai referensi visual yang penting dalam penelitian ini.

B. Preprocessing

Setelah *dataset* terkumpul sebanyak 500 citra, terdapat beberapa masalah pada *dataset* yaitu *size* atau ukuran dari gambar atau citra rumah adat yang berbeda-beda, baik dalam tinggi maupun lebar. Selain itu, terdapat *noise* atau derau pada citra yaitu objek lain yang tidak diperlukan seperti objek orang hewan, batu kubur, dan lain-lain. sehingga perlu dilakukan proses penghilangan *noise*. Proses penghilangan derau pada cira menggunakan aplikasi *snapedit*.

Setelah itu untuk menyelesaikan persoalan variasi ukuran dan kualitas citra melalui proses normalisasi dan *resizing* menjadi 256x 256 untuk seluruh gambar. Proses selanjutnya adalah pelabelan data dan diketahui bahwa terdapat 5 jenis rumah adat yang akan dikelompokkan sesuai kelas tertentu, yaitu kelas 0 untuk RA Bajawa, kelas 1 RA Ende, kelas 2 RA Pulau Timur, kelas 3 untuk RA Sumba, dan kelas 4 untuk RA Waerebo.

Terakhir, untuk memastikan kualitas pembelajaran model, dilakukan pemisahan *dataset* menjadi data latih, data validasi dan data uji. Pemisahan *dataset* ini peneliti menggunakan *train_test_split* dengan pembagian data latih sebesar 0.7

sehingga mendapatkan 350 citra untuk data latih, 105 untuk data validasi dan 45 untuk data uji (Tabel 1).

Tabel 1. *Preprocessing* Penggunaan Model

Model	Resize Image	Pembagian Dataset	Learning Rate
CNN	256 x 256	70	0,0001
CapsNet	256 x 256	70	0,0001

C. Pengukuran Model

Pengukuran kinerja model sangat penting untuk mengetahui seberapa baik model dapat mengklasifikasikan data ke dalam kelas-kelas tertentu. Pada penelitian ini, kinerja model dinilai dengan *confusion matrix*. *Confusion matrix* adalah alat evaluasi pada *machine learning* yang berisi informasi perbandingan antara kelas prediksi model dengan kelas aktual data [25]. *Confusion matrix* terdiri dari empat (4) elemen yaitu *True Positif* (TP) menunjukkan model dapat memprediksi secara benar untuk kelas data aktual positif atau milik kelas tersebut, *True Negative* (TN) menunjukkan model memprediksi secara benar untuk kelas data aktual negatif, namun dalam klasifikasi multi kelas TN diabaikan karena data dari sebuah kelas tidak memiliki hubungan dengan kelas yang lain. *False Positive* (FP) menunjukkan model memprediksi secara salah untuk kelas data aktual positif, dan *False Negative* (FN) menunjukkan model memprediksi secara salah untuk kelas data aktual negatif.

Dalam klasifikasi multi kelas yang menjadi kasus klasifikasi dalam penelitian ini, *confusion matrix* yang digunakan adalah ukuran 5x5 karena data aktual yang digunakan terdiri dari 5 kelas yaitu RA Bajawa sebagai kelas 0, RA Ende sebagai kelas 1, RA Pulau Timur sebagai kelas 2, RA Sumba sebagai kelas 3, dan RA Waerebo sebagai kelas 4.

Dari *Confusion matrix* dapat dihitung nilai akurasi, presisi, *recall*, maupun *F1-score* untuk setiap kelas. Dalam penelitian ini yang digunakan adalah evaluasi akurasi dari model. Akurasi adalah indikator yang digunakan untuk mengevaluasi

seberapa akurat model mengklasifikasikan dengan benar. Dimana nilai akurasi diperoleh melalui rumus (1).

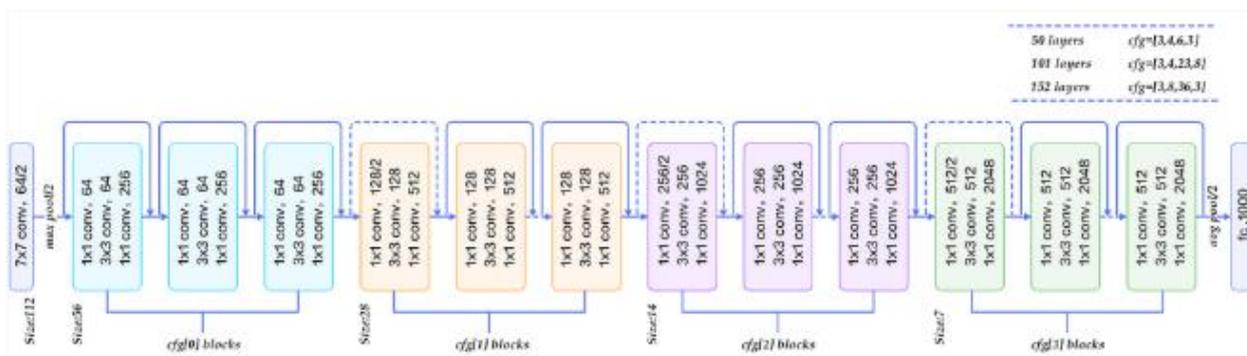
$$accuracy = \frac{TP}{TP+TN+FP+FN} \tag{1}$$

Dalam kasus ini menentukan nilai akurasi yaitu jumlah prediksi benar dibagi total seluruh data yang digunakan.

D. Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNN) adalah jenis jaringan saraf yang digunakan secara luas dalam pengolahan citra dan pengenalan pola. Dalam pengujian ini, peneliti menggunakan model *ResNet50V2*, yang merupakan varian dari *ResNet* (*Residual Network*) yang telah terbukti sangat efektif dalam berbagai tugas pengolahan citra. Arsitektur *ResNet50V2* memanfaatkan lapisan konvolusi dalam yang mendalam dan koneksi-koneksi residu yang memungkinkannya untuk mengatasi masalah hilangnya gradien dan pelatihan model yang lebih mendalam (Gambar 3). Setelah lapisan konvolusi, terdapat lapisan *Global Average Pooling* (GAP) yang mengurangi dimensi fitur, diikuti oleh beberapa lapisan tersembunyi (*Dense*) yang menghasilkan *output* klasifikasi. Pemilihan *ResNet50V2* sebagai model dasar dalam penelitian ini didasarkan pada kemampuannya yang terbukti dalam mengekstrak fitur-fitur penting dari citra secara efektif dan mengatasi masalah pelatihan model yang dalam [26].

Pada model *ResNet50V2*, lapisan GAP diimplementasikan untuk mengurangi dimensi fitur sehingga meningkatkan efisiensi pengolahan dan menghasilkan representasi fitur yang lebih ringkas [27]. Lapisan GAP ini diikuti oleh beberapa lapisan tersembunyi (*Dense*) yang bertanggung jawab untuk menghasilkan *output* klasifikasi. Dengan menggunakan *ResNet50V2*, penelitian ini bertujuan untuk menguji sejauh mana model ini dapat mengklasifikasikan citra rumah adat NTT dengan akurasi yang tinggi dan efisien.



Gambar 3. Arsitektur Model CNN [18]

ResNet50V2 adalah salah satu arsitektur jaringan saraf konvolusi CNN yang merupakan bagian dari keluarga *Residual Networks* atau *ResNet*. Arsitektur ini digunakan untuk pengolahan citra, termasuk klasifikasi gambar. Pertama, lapisan konvolusi berfungsi sebagai pengambil fitur awal yang digunakan untuk menangkap detail penting dari

gambar. Lapisan ini menggunakan filter atau inti untuk mendeteksi berbagai fitur seperti tepi, tekstur, atau bentuk di dalam gambar.

Kedua, yang membedakan *ResNet* dari arsitektur jaringan saraf lainnya adalah penggunaan blok sisa atau "lapisan sisa". Blok-blok ini memungkinkan informasi dari lapisan

sebelumnya untuk "melompati" atau langsung diteruskan ke lapisan berikutnya. Ini membantu mengatasi masalah pelatihan yang lebih dalam dengan mencegah hilangnya gradien. Setiap blok sisa memiliki dua jalur, yaitu jalur langsung yang mengirimkan informasi langsung dari lapisan sebelumnya, dan jalur utama dimana transformasi konvolusi diterapkan pada masukan.

Terakhir, setelah melalui serangkaian lapisan konvolusi, *ResNet50V2* menggunakan lapisan pengambilan rata-rata secara global untuk mengurangi dimensi fitur dengan menghitung nilai rata-rata dari setiap saluran di seluruh wilayah gambar. Kemudian terdapat lapisan terhubung sepenuhnya yang bertugas menghasilkan keluaran klasifikasi. Lapisan-lapisan ini memanfaatkan fungsi aktivasi seperti *ReLU* untuk memperkenalkan non-linearitas ke dalam model dan menggunakan fungsi kerugian seperti silang entropi kategorial selama pelatihan, yang dijalankan dengan bantuan algoritma optimasi seperti *Gradient Descent Stoastik (SGD)* atau Adam untuk memperbarui bobot dalam jaringan. Semua fungsi ini bekerja sama untuk membantu *ResNet50V2* mengidentifikasi dan mengklasifikasikan gambar dengan akurasi tinggi.

Secara keseluruhan, *ResNet50V2* adalah jaringan yang sangat dalam dan kuat yang mampu mengekstrak fitur-fitur penting dari citra dan menghasilkan prediksi kelas yang akurat. Penggunaan *residual blocks* memungkinkan pelatihan jaringan yang lebih dalam tanpa menghadapi masalah hilangnya gradien, yang seringkali menjadi masalah dalam jaringan yang dalam. Dengan cara kerja ini, *ResNet50V2* telah terbukti efektif dalam berbagai tugas pengolahan citra.

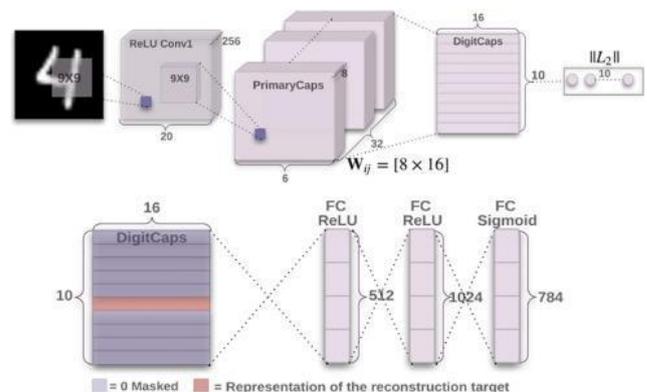
D. Capsule Network

Capsule Network (CapsNet) adalah inovasi terbaru dalam jaringan saraf buatan yang memperkenalkan konsep "kapsul" atau kapsul sebagai unit pemrosesan utama. Salah satu varian *CapsNet* yang digunakan dalam penelitian ini adalah model *CapsNet-COVID19*, yang dikembangkan khusus untuk tugas klasifikasi citra CT paru-paru dalam konteks *COVID-19* [21]. *CapsNet* memiliki kemampuan untuk mengatasi beberapa masalah yang dihadapi oleh CNN, seperti masalah ketidakjelasan dalam citra. Hal ini dicapai melalui struktur hierarki yang lebih baik dalam pengenalan fitur-fitur penting dalam citra. Dalam *CapsNet*, kapsul-kapsul ini dapat bekerja bersama untuk mendeteksi fitur-fitur dalam berbagai posisi dan orientasi, menghasilkan representasi fitur yang lebih kuat dan konsisten.

Dalam model *CapsNet-COVID19*, *CapsNet* digunakan untuk mengklasifikasikan citra CT paru-paru sebagai positif atau negatif *COVID-19* [21]. Penggunaan *CapsNet* dalam konteks ini bertujuan untuk memanfaatkan keunggulan *CapsNet* dalam pengenalan fitur-fitur yang mungkin memiliki variasi posisi atau rotasi pada citra CT. Dengan demikian, penelitian ini akan menguji sejauh mana *CapsNet-COVID19* dapat mengklasifikasikan citra CT paru-paru dengan tingkat akurasi yang tinggi, serta apakah model ini memiliki keunggulan dibandingkan dengan metode tradisional seperti CNN dalam tugas klasifikasi citra medis.

Penelitian ini akan menguji seberapa baik metode *Capsule Network (CapsNet)* dapat mengklasifikasikan citra rumah adat, dan apakah model ini lebih unggul dibanding metode CNN dalam tugas yang sama. Kinerja kedua model akan dievaluasi berdasarkan dua ukuran yaitu akurasi dan *loss*, untuk melihat mana yang lebih baik dalam mengenali dan mengklasifikasikan citra rumah adat dari berbagai suku di Nusa Tenggara Timur.

CapsNet memperkenalkan konsep "kapsul" sebagai unit pemrosesan utama yang berbeda dari lapisan konvensional dalam CNN. Kapsul dalam *CapsNet* didesain untuk mendeteksi dan merepresentasikan fitur-fitur dalam citra dengan cara yang lebih hierarkis, mengatasi masalah ketidakjelasan dan variasi posisi dalam citra. Kapsul-kapsul ini bekerja bersama untuk mengenali fitur-fitur dalam berbagai orientasi dan posisi, sehingga menghasilkan representasi fitur yang lebih kuat dan konsisten seperti pada Gambar 4. Hal ini memungkinkan *CapsNet* untuk mengatasi beberapa tantangan yang dihadapi oleh CNN, termasuk dalam pengenalan citra dengan ketidakpastian dalam struktur objek atau objek dengan variasi orientasi yang tinggi. Dengan demikian, *CapsNet* menjadi alat yang menjanjikan dalam pengolahan citra yang membutuhkan pemahaman fitur yang lebih kontekstual dan lebih kuat.



Gambar 4. Model Arsitektur *CapsNet* [28]

III. HASIL DAN PEMBAHASAN

1. Pelatihan Model CNN dan *CapsNet*

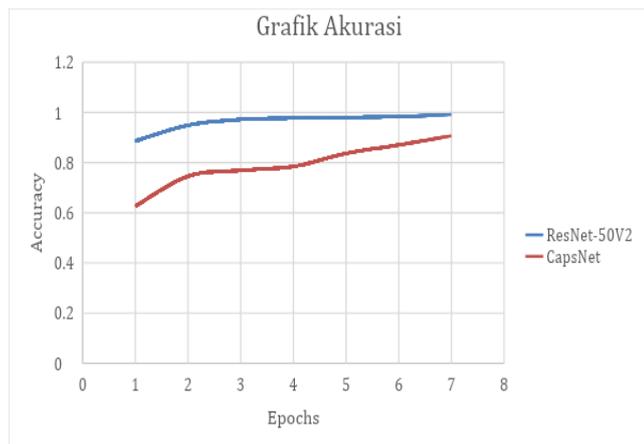
Pelatihan model dilakukan dengan menggunakan beberapa parameter, yaitu *batch size* 15, *learning rate* 0,0001, serta menggunakan *optimizer Adam*, dan *loss function* menggunakan *categorical_crossentropy*. Pelatihan model dilakukan menggunakan *Google Colab* dengan spesifikasi 12GB RAM dan *GPU T4* dengan kapasitas 15GB. *Framework* yang digunakan adalah *numpy*, *pandas*, *matplotlib*, *pyplot*, *scikit-learn*, *Tensorflow*, dan *Keras*. Terdapat beberapa perbedaan dalam pelatihan model, pertama, arsitektur CNN *ResNet50V2* hanya membutuhkan 7 *epoch* untuk mencapai hasil yang signifikan, sedangkan model *CapsNet* 300 *epoch*. Selain itu, terdapat perbedaan dalam jumlah parameter antara kedua model ini, dimana CNN dengan arsitektur *ResNet50V2* memiliki jumlah parameter sebanyak 24,090,629, sedangkan

CapsNet memiliki 390,691,285 parameter, hal ini menunjukkan kompleksitas yang lebih tinggi pada model *CapsNet*.

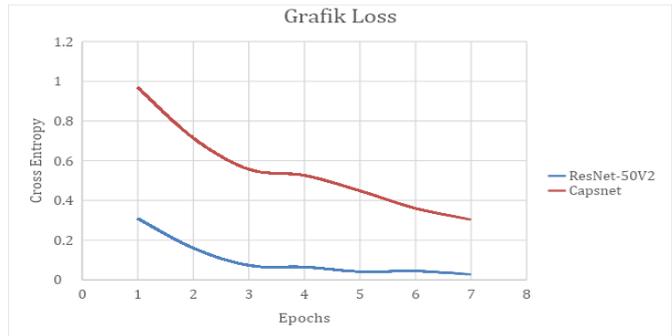
Pada model CNN, terdapat empat lapisan/layer yaitu *ResNet50v2*, *GlobalAveragePooling2D*, *dense_2*, dan *dense_3*. *ResNet50v2* digunakan sebagai model dasar karena mengacu pada arsitektur *ResNet*. *GlobalAveragePooling2D* berguna untuk menghitung rata-rata nilai seluruh *channel* dari lapisan *ResNet50v2*. *dense_2* memproses fitur yang telah diekstrak oleh lapisan sebelumnya untuk menghasilkan representasi yang lebih padat. Sedangkan pada model *CapsNet*, terdapat *input_1* sebagai input data masuk, *Conv2D* untuk mengekstrak fitur dari data masukan, *MaxPooling2D* untuk mengurangi dimensi spasial dengan mengambil nilai maksimum, *Flatten* menghubungkan lapisan konvolusi ke lapisan *dense*, dan terakhir *Dense* untuk melakukan klasifikasi. *Input_1* berfungsi sebagai titik awal masuknya data ke dalam model. *Conv2D* digunakan untuk mengekstraksi fitur dari data masukan. *MaxPooling2D* berguna untuk mengurangi ukuran data dengan cara mengambil nilai maksimum. *Flatten* menyatukan lapisan konvolusi sehingga dapat dihubungkan ke lapisan *Dense*. Dan lapisan *Dense* terakhir digunakan untuk melakukan klasifikasi.

Tabel 2. Perbandingan Hasil CNN dan *CapsNet*

	Jumlah Parameter	Akurasi	Loss
<i>ResNet50V2</i>	24.090.629	98%	0,72
<i>CapsNet</i>	390.691.285	72%	1,37



Gambar 5. Nilai Akurasi CNN dan *CapsNet*



Gambar 6. Nilai Loss CNN dan *CapsNet*

Dari Tabel 2 di atas, menggambarkan parameter, *epoch*, *batch size*, dan *learning rate* antara dua algoritma *deep learning* yang digunakan dalam penelitian ini, yaitu CNN dan *CapsNet*.

Pada Gambar 5 menunjukkan grafik akurasi dimana jika akurasi semakin tinggi maka titik akurasi semakin menuju ke atas. Pada grafik tersebut di lihat bahwa nilai akurasi untuk CNN lebih tinggi dibandingkan dengan grafik *CapsNet*. Nilai akurasi semakin meningkat ke atas menandakan model semakin baik dalam melakukan prediksi dengan benar. Sedangkan pada Gambar 6 menunjukkan jika grafik *loss* semakin menurun ke bawah, maka model semakin baik mengurangi deviasi antara hasil prediksi dan nilai sebenarnya. Pada grafik tersebut di lihat bahwa nilai *loss* untuk CNN lebih rendah dibandingkan dengan grafik *CapsNet*.

Untuk CNN dan *CapsNet*, *learning rate* yang digunakan sama yaitu sebesar 0,0001 dengan jumlah *epoch* sebanyak 300 dan ukuran *batch size* sebesar 15. Perbandingan ini merefleksikan adanya perbedaan dalam konfigurasi pelatihan antara kedua algoritma, dimana hal ini dapat memengaruhi kinerja dan kecepatan konvergensi masing-masing model. Tujuan dari perbedaan parameter ini adalah untuk mendapatkan hasil pelatihan terbaik dari setiap algoritma, mengingat karakteristik pelatihan yang berbeda antara CNN dan *CapsNet*.

2. Analisis Hasil Perbandingan

Analisis hasil penelitian mengungkapkan adanya perbedaan signifikan dalam kinerja antara Jaringan Saraf Konvolusi CNN dan *CapsNet* pada *dataset* citra rumah adat NTT. Pertama, perbandingan akurasi menunjukkan bahwa CNN (*ResNet50V2*) secara konsisten memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan *CapsNet*. Hal ini mengindikasikan bahwa dalam tugas klasifikasi citra rumah adat, arsitektur CNN lebih efektif dalam mengenali dan mengklasifikasikan citra secara akurat.

Meskipun *CapsNet* menunjukkan potensi dalam pengenalan fitur citra yang memiliki variasi posisi atau rotasi, model ini masih perlu ditingkatkan untuk *dataset* ini. Analisis lebih lanjut mengaitkan hasil ini dengan penelitian sebelumnya yang menunjukkan *CapsNet* lebih akurat dalam pengenalan angka [15]. Hal ini mengindikasikan bahwa *CapsNet* mungkin lebih cocok untuk tugas-tugas klasifikasi fitur-fitur lebih terstruktur. Namun, pengenalan rumah adat menjadi tantangan yang menarik dalam penelitian ini. Karena

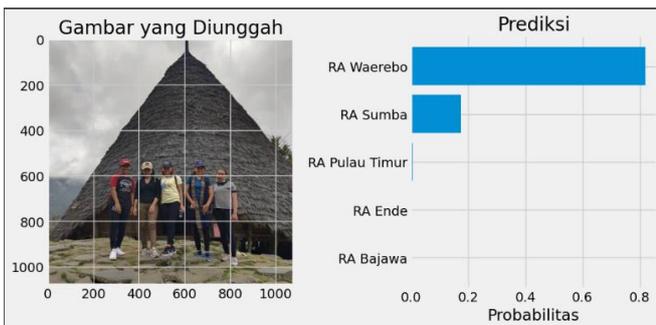
rumah adat memiliki desain, struktur, dan elemen budaya yang hampir sama, pengenalan dan klasifikasi rumah adat menjadi lebih sulit dibandingkan dengan pengenalan angka atau objek lain yang lebih terstruktur.

3. Perbandingan Hasil Klasifikasi

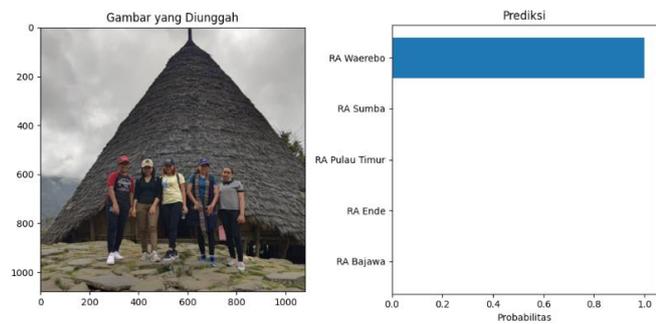
Dalam menganalisis hasil klasifikasi antara CNN dan CapsNet, terdapat temuan menarik yang patut diperhatikan. Meskipun secara akurasi CNN unggul, terdapat aspek di mana CapsNet menunjukkan keunggulan. Khususnya, dalam konteks pengenalan rumah adat NTT, CapsNet mampu memprediksi dengan lebih baik. Hal ini menunjukkan bahwa meskipun akurasi umum CapsNet masih perlu ditingkatkan, model ini memiliki potensi dalam pengenalan citra dengan karakteristik khusus [29]. Kinerja model klasifikasi sangat bergantung pada karakteristik dataset dan jenis tugas yang diberikan. Umumnya CNN unggul dalam pengenalan gambar dengan fitur ekstraksi hierarkisnya. Namun untuk dataset rumah adat ini, CapsNet mungkin lebih baik memahami hubungan hierarkis antar bagian dan keseluruhan rumah adat. Ini karena CapsNet lebih handal mengatasi masalah bagian-keseluruhan dibanding CNN. Sementara itu, CNN mungkin kesulitan jika variasi posisi atau orientasi objek bergaya rumah adat menjadi faktor penting dalam dataset. Hal ini dapat dilihat dengan lebih jelas pada Gambar 7 dan Gambar 8, di mana perbandingan hasil klasifikasi kedua algoritma tersebut.

dengan baik menggunakan CNN dengan *evidence* 0,8 untuk kelas waerebo (RA Waerebo) dan 0,2 untuk kelas rumah adat Sumba (RA Sumba). Hal tersebut disebabkan kemiripan bentuk atap kedua rumah adat tersebut. Namun CNN dapat disimpulkan mampu melakukan prediksi atau mengklasifikasikan rumah adat Waerebo ke kelas Waerebo (RA Waerebo) dengan baik. Gambar 8 menunjukkan hasil prediksi yang dilakukan oleh algoritma CapsNet, dimana data gambar rumah adat, yaitu aktual waerebo diklasifikasikan dengan benar sebagai rumah adat Waerebo (RA Waerebo) dengan hasil prediksi atau hasil klasifikasi tepat dengan skor 1,0.

Selain itu juga terdapat hasil prediksi atau hasil klasifikasi yang salah dari kedua algoritma tersebut. Gambar 9 menunjukkan hasil prediksi atau klasifikasi yang salah dari algoritma CNN, dimana data aktual rumah adat Sao Ata Mosa Lakitana (RA Sumba) diprediksi atau diklasifikasikan ke kelas rumah ada Sao Ria (RA Ende). Gambar 10 menunjukkan hasil prediksi atau klasifikasi yang salah dari algoritma CapsNet, dimana data aktual rumah adat Mbaru Niang suku Waerebo (RA Waerebo) diprediksi atau diklasifikasikan ke kelas rumah adat Sao Ata Mosa Lakitana (RA Sumba).



Gambar 7. Prediksi CNN (ResNet50V2)



Gambar 8. Prediksi Capsule Network

Gambar 7 dan Gambar 8 merupakan contoh pengujian terhadap data rumah adat yang tidak termasuk dalam data pelatihan maupun data pengujian. Gambar 7 menunjukkan data aktual rumah adat Mbaru Niang suku Waerebo diprediksi



Gambar 9. Prediksi Salah Algoritma CNN



Gambar 10. Prediksi Salah Algoritma CapsNet

IV. KESIMPULAN

Dalam kasus pengenalan citra rumah adat NTT, penelitian ini menemukan hal-hal penting terkait kinerja CNN dan *CapsNet*, yaitu akurasi CNN lebih tinggi dengan mencapai 98% dan nilai loss yang lebih rendah yaitu 0,72%, namun *CapsNet* menunjukkan potensi dalam mengklasifikasi citra yang mungkin memiliki variasi posisi atau rotasi yang kompleks. Khususnya, algoritma *CapsNet* dapat memprediksi dengan lebih baik dalam kasus-kasus tertentu.

Dalam proses pelatihan model menunjukkan bahwa *CapsNet* membutuhkan pelatihan lebih lama dan memiliki kompleksitas model yang tinggi. Oleh karena itu, pengembangan *CapsNet* untuk meningkatkan akurasi model perlu dilakukan dalam penelitian yang lebih lanjut. Temuan ini menekankan pentingnya memahami perbedaan CNN dan *CapsNet* dalam pengenalan citra. Perbandingan klasifikasi benar dan klasifikasi salah antara kedua algoritma menunjukkan keunggulan masing-masing dalam situasi tertentu.

Oleh karena itu, perbandingan kedua algoritma ini memberikan informasi mengenai kinerja CNN dan *CapsNet* dalam pengenalan citra rumah adat NTT. CNN menjadi pilihan yang lebih baik, dikarenakan CNN memiliki akurasi yang tinggi, sementara *CapsNet* unggul dalam pengenalan citra dengan variasi posisi dan rotasi yang kompleks.

Dalam hal pengaturan *learning rate*, baik untuk CNN maupun *CapsNet*, dalam penelitian ini menggunakan nilai *learning rate* sebesar 0,0001. Penggunaan nilai ini bertujuan untuk melatih model dengan perubahan bobot yang sangat kecil pada setiap iterasi. Hal ini sangat penting saat melakukan *fine tuning* pada model yang sudah dilatih sebelumnya. Dengan menggunakan *learning rate* yang rendah, dapat menjaga kestabilan dan konsistensi pelatihan model saat menyesuaikan dengan *dataset* rumah adat Nusa Tenggara Timur (NTT).

Berdasarkan hasil akurasi klasifikasi dan kesesuaian dengan *dataset* yang digunakan, maka penelitian ke depan yang dilakukan adalah mengimplementasikan model CNN dengan mengembangkan aplikasi berbasis *mobile* untuk mengidentifikasi rumah adat, sehingga bisa digunakan oleh masyarakat umum, pelajar, maupun wisatawan, dalam identifikasi rumah adat menggunakan perangkat *mobile* dari fitur kamera. Selain itu, pengembangan jaringan *Capsule Networks* perlu dilakukan untuk menaikkan nilai akurasi klasifikasi dengan jumlah *dataset* yang lebih banyak dan penggunaan *hyperparameter tuning*.

REFERENSI

[1] T. Abdulghani and B. P. Sati, "Pengenalan Rumah Adat Indonesia Menggunakan Teknologi Augmented Reality Dengan Metode Marker Based Tracking Sebagai Media Pembelajaran," *Media J. Inform.*, vol. 11, no. 1, p. 43, 2020, doi: 10.35194/mji.v11i1.770.

[2] N. K. Sutiari, I. K. G. Darma Putra, and I. M. Sunia Raharja, "Aplikasi Pengenalan Rumah Adat Indonesia Berbasis Augmented Reality," *J. Ilm. Merpati (Menara*

Penelit. Akad. Teknol. Informasi), vol. 6, no. 2, p. 108, 2018, doi: 10.24843/jim.2018.v06.i02.p05.

- [3] Q. Aini, N. Lutfiani, H. Kusumah, and M. S. Zahran, "Deteksi dan Pengenalan Objek Dengan Model Machine Learning: Model Yolo," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 6, no. 2, p. 192, 2021, doi: 10.24114/cess.v6i2.25840.
- [4] V. Ayumi, "Pengenalan Objek Bunga Berbasis Deep Learning Menggunakan Model Resnet50 dan MobileNet-v2," *JSAI (Journal Sci. Appl. Informatics)*, vol. 5, no. 3, pp. 188–193, 2022, doi: 10.36085/jsai.v5i3.3613.
- [5] S. H. Abdullah, R. Magdalena, and R. Y. N. Fu'adah, "Klasifikasi Diabetic Retinopathy Berbasis Pengolahan Citra Fundus Dan Deep Learning," *J. Electr. Syst. Control Eng.*, vol. 5, no. 2, pp. 84–90, 2022, doi: 10.31289/jesce.v5i2.5659.
- [6] A. J. Rozaqi, A. Sunyoto, and M. rudyanto Arief, "Deteksi Penyakit Pada Daun Kentang Menggunakan Pengolahan Citra dengan Metode Convolutional Neural Network," *Creat. Inf. Technol. J.*, vol. 8, no. 1, p. 22, 2021, doi: 10.24076/citec.2021v8i1.263.
- [7] A. Sani and S. Rahmadinni, "Deteksi Gestur Tangan Berbasis Pengolahan Citra," *J. Rekayasa Elektr.*, vol. 18, no. 2, pp. 115–124, 2022, doi: 10.17529/jre.v18i2.25147.
- [8] V. Wahyuningtyas, "Implementasi Ekstraksi Fitur untuk Klasifikasi Suara Urban Menggunakan Deep Learning," vol. 3, no. 1, p. 10, 2021.
- [9] N. Azis, H. Herwanto, and F. Ramadhani, "Implementasi Speech Recognition Pada Aplikasi E-Prescribing Menggunakan Algoritme Convolutional Neural Network," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 460, 2021, doi: 10.30865/mib.v5i2.2841.
- [10] Rendi Nurcahyo and Mohammad Iqbal, "Pengenalan Emosi Pembicara Menggunakan Convolutional Neural Networks," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 115–122, 2022, doi: 10.29207/resti.v6i1.3726.
- [11] I. M. Dendi Maysanjaya, "Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan Convolutional Neural Network (Classification of Pneumonia Based on Lung X-rays Images using Convolutional Neural Network)," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 9, no. 2, p. 190, 2020, [Online]. Available: <https://garuda.kemdikbud.go.id/documents/detail/2807288>
- [12] Y. F. Riti and S. S. Tandjung, "Klasifikasi Covid-19 Pada Citra CT Scans Paru-Paru Menggunakan Metode Convolution Neural Network," *Progresif J. Ilm. Komput.*, vol. 18, no. 1, p. 91, 2022, doi: 10.35889/progresif.v18i1.784.
- [13] D. Gunawan and H. Setiawan, "Convolutional Neural Network dalam Citra Medis," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 376–390, 2022, doi: 10.24002/konstelasi.v2i2.5367.
- [14] M. Rahimzadeh, A. Attar, and S. M. Sakhaei, "A fully automated deep learning-based network for detecting

- COVID-19 from a new and large lung CT scan dataset,” *Biomed. Signal Process. Control*, vol. 68, no. March, p. 102588, 2021, doi: 10.1016/j.bspc.2021.102588.
- [15] K. Adu, Y. Yu, J. Cai, K. Owusu-Agyemang, B. A. Twumasi, and X. Wang, “DHS-CapsNet: Dual horizontal squash Capsule Networks for lung and colon cancer classification from whole slide histopathological images,” *Int. J. Imaging Syst. Technol.*, vol. 31, no. 4, pp. 2075–2092, 2021, doi: 10.1002/ima.22569.
- [16] M. Harahap, Em Manuel Laia, Lilis Suryani Sitanggang, Melda Sinaga, Daniel Franci Sihombing, and Amir Mahmud Husein, “Deteksi Penyakit Covid-19 Pada Citra X-Ray Dengan Pendekatan Convolutional Neural Network (CNN),” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 70–77, 2022, doi: 10.29207/resti.v6i1.3373.
- [17] A. S. Wiratama, M. Rifqi, and S. Maesaroh, “Efektivitas Transfer Learning Dalam Pendeteksian Penyakit Pneumonia Melalui Citra X-Ray Paru Manusia,” *J. Ilm. Sains dan Teknol.*, vol. 7, no. 1, pp. 43–52, 2023, doi: 10.47080/saintek.v7i1.2551.
- [18] M. V. Cruz, A. Namburu, S. Chakkaravarthy, M. Pittendreigh, and S. C. Satapathy, “Skin cancer classification using convolutional capsule network (CapsNet),” *J. Sci. Ind. Res. (India)*, vol. 79, no. 11, pp. 994–1001, 2020.
- [19] K. Aljohani and T. Turki, “Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks,” *Ai*, vol. 3, no. 2, pp. 512–525, 2022, doi: 10.3390/ai3020029.
- [20] X. Jiang *et al.*, “Hyperspectral image classification with CapsNet and markov random fields,” *IEEE Access*, vol. 8, pp. 191956–191968, 2020, doi: 10.1109/ACCESS.2020.3029174.
- [21] X. Q. Zhang, G. Y. Wang, and S. G. Zhao, “CapsNet-COVID19: Lung CT image classification method based on CapsNet model,” *Math. Biosci. Eng.*, vol. 19, no. 5, pp. 5055–5074, 2022, doi: 10.3934/mbe.2022236.
- [22] A. Muhammad, M. A. Arserim, and Ö. Türk, “Compare the classification performances of convolutional neural networks and Capsule Networks on the Coswara dataset,” vol. 2, pp. 265–271, 2023, doi: 10.24012/dumf.1270429.
- [23] U. Kulkarni, S. M. Meena, S. V. Gurlahosur, and U. Mudengudi, “Classification of cultural heritage sites using transfer learning,” *Proc. - 2019 IEEE 5th Int. Conf. Multimed. Big Data, BigMM 2019*, pp. 391–397, 2019, doi: 10.1109/BigMM.2019.00020.
- [24] M. Cosovic and R. Jankovic, “CNN Classification of the Cultural Heritage Images,” *2020 19th Int. Symp. INFOTEH-JAHORINA, INFOTEH 2020 - Proc.*, no. March, pp. 18–20, 2020, doi: 10.1109/INFOTEH48170.2020.9066300.
- [25] G. Zeng, “On the confusion matrix in credit scoring and its analytical properties,” *Commun. Stat. - Theory Methods*, vol. 49, no. 9, pp. 2080–2093, 2020, doi: 10.1080/03610926.2019.1568485.
- [26] A. W. Setiawan, “Perbandingan Arsitektur Convolutional Neural Network Pada Klasifikasi Pneumonia, COVID-19, Lung Opacity, dan Normal Menggunakan Citra Sinar-X Thoraks,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 7, p. 1563, 2022, doi: 10.25126/jtiik.2022976742.
- [27] N. Makanapura, C. Sujatha, P. R. Patil, and P. Desai, “Classification of plant seedlings using deep Convolutional Neural Network architectures,” *J. Phys. Conf. Ser.*, vol. 2161, no. 1, 2022, doi: 10.1088/1742-6596/2161/1/012006.
- [28] M. Kwabena Patrick, A. Felix Adekoya, A. Abra Mighty, and B. Y. Edward, “Capsule Networks – A survey,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 1, pp. 1295–1310, 2022, doi: 10.1016/j.jksuci.2019.09.014.
- [29] E. Xi, S. Bing, and Y. Jin, “Capsule Network Performance on Complex Data,” vol. 10707, no. Fall, pp. 1–7, 2017, [Online]. Available: <http://arxiv.org/abs/1712.03480>