

PERBANDINGAN VARIAN MODEL *EFFICIENTNETV2* PADA CITRA HISTOLOGI *OSTEOSARCOMA*

Youllia Indrawaty Nurhasanah¹, Rifasya Ayu Nurfayza²

^{1,2}Institut Teknologi Nasional, Bandung
Email: ¹youllia1975@gmail.com, ²nurfayza.rr@gmail.com
*Penulis Korespondensi

(Naskah masuk: 20 Oktober 2024, diterima untuk diterbitkan: 29 Oktober 2025)

Abstrak

Osteosarcoma merupakan jenis kanker tulang ganas yang menyerang ujung tulang panjang dan berpotensi menyebar ke organ lain (metastasis). Diagnosis dini berperan penting untuk mendukung hasil pengobatan yang optimal. Penelitian ini menggunakan *Convolutional Neural Network* (CNN) untuk menganalisis dan mengklasifikasikan citra histologi tulang *Osteosarcoma*, dengan membandingkan kinerja tiga varian model *EfficientNetV2* (S, M, dan L). Dataset yang digunakan adalah citra histologi *Osteosarcoma* yang telah didigitalisasi dan dataset ini diproses melalui tahap *preprocessing*, augmentasi, serta training menggunakan ketiga model, lalu diproses pula melalui konfigurasi hyperparameter. Evaluasi kinerja model dilakukan berdasarkan akurasi, Presisi, Recall, dan *F1-Score*. Hasil penelitian menunjukkan bahwa *EfficientNetV2-S* mencapai akurasi tertinggi sebesar **88,86%** dengan efisiensi yang lebih baik, sedangkan *EfficientNetV2-M* memiliki stabilitas klasifikasi yang lebih baik dengan *F1-Score* yang lebih konsisten dengan akurasi sebesar 88,80%. Sementara itu, *EfficientNetV2-L* menunjukkan hasil akurasi yang kompetitif tetapi memerlukan sumber daya komputasi yang lebih besar. Analisis lebih lanjut menunjukkan bahwa pemilihan model tidak hanya bergantung pada akurasi, tetapi juga mempertimbangkan ukuran model dan kebutuhan komputasi. Hasil penelitian ini menunjukkan bahwa *EfficientNetV2-S* merupakan pilihan optimal berdasarkan akurasi dan efisiensi, sedangkan *EfficientNetV2-M* lebih unggul dalam stabilitas klasifikasi. Hasil penelitian ini dapat menjadi referensi dalam pengembangan sistem berbasis deep learning untuk diagnosis kanker tulang di masa depan.

Kata kunci: Akurasi, Diagnosis awal, *EfficientNetV2*, histologi tulang, *Osteosarcoma*, klasifikasi

COMPARISON OF *EFFICIENTNETV2* MODEL VARIANTS ON *OSTEOSARCOMA* HISTOLOGY IMAGES

Abstract

Osteosarcoma is a malignant bone cancer that primarily affects the ends of long bones and has a high potential for metastasis to other organs. Early diagnosis is crucial to improving treatment outcomes and patient prognosis. This study employs a *Convolutional Neural Network* (CNN) to analyze and classify histological images of *Osteosarcoma* by comparing the performance of three variants of the *EfficientNetV2* model (S, M, and L). The dataset that used in this study is a digitized *Osteosarcoma* histology images, and will be processed through the *preprocessing*, data augmentation, and model training using the three different *EfficientNetV2* variants. Additionally, hyperparameter tuning is performed to optimize model performance. The evaluation of model performance is conducted based on accuracy, precision, recall, and *F1-score*. The results showed that *EfficientNetV2-S* achieved the highest accuracy of 88.86% with better efficiency, while *EfficientNetV2-S* had better classification stability with a more consistent *F1-Score* with an accuracy of 88.80%. Meanwhile, *EfficientNetV2-L* showed competitive accuracy results but required more computational resources. Further analysis reveals that model selection not only depends on accuracy, but also considers model size and computational requirements. The results show that *EfficientNetV2-S* is the optimal choice based on accuracy and efficiency, while *EfficientNetV2-M* is the optimal choice in classification stability. The results of this study can serve as a reference in the development of deep learning-based systems for bone cancer diagnosis in the future.

Keywords: Accuracy, Early Diagnosis, *EfficientNetV2*, Bone Histology, *Osteosarcoma*, Classification.

1. PENDAHULUAN

Convolutional Neural Network (CNN) merupakan salah satu jenis algoritma pembelajaran

deep learning yang dapat memproses, mengenali, mengklasifikasikan, dan mengolah citra atau gambar (ANHAR & PUTRA, 2023). CNN telah diterapkan di berbagai bidang, salah satunya diterapkan pada

bidang medis seperti untuk pengolahan citra *Magnetic Resonance Imaging* (MRI), *Computed Tomography* (CT), *ultrasonografi* (USG), sinar-X, histologi, dan *Positron Emission Tomography* (PET), dan masing-masing citra tersebut memiliki fungsi spesifik dalam diagnosis medis (Gunawan & Setiawan, 2022). Penerapan yang dilakukan adalah seperti klasifikasi, deteksi, diagnosis awal penyakit seperti penyakit kanker.

Pada penelitian ini, CNN digunakan untuk melakukan klasifikasi dan diagnosis awal penyakit *Osteosarcoma*. Penyakit ini merupakan kanker tulang yang berkembang pada tulang panjang manusia, seperti pada tulang paha, lengan atas, dan lutut. Kasus terjadinya *Osteosarcoma* dalam satu tahun mencapai 4 – 5 kasus per satu juta penduduk menurut *World Health Organization* (WHO). Penyebaran penyakit ini dapat tersebar ke organ lain dengan cepat (Anisuzzaman et al., 2021). Untuk mengurangi resiko penyebaran ke organ lain (metastasis) sangat penting dilakukannya diagnosis awal untuk penyakit ini. Diagnosis awal dapat dilakukan dengan prosedur biopsi, yang merupakan prosedur pengambilan sampel kecil jaringan tubuh manusia untuk diperiksa di laboratorium dengan tujuan mendeteksi keberadaan suatu penyakit. Hasil dari prosedur ini adalah citra histologi (MAWARDI et al., 2024). Citra histologi ini dilengkapi dengan pewarnaan *Hematoxylin & Eosin* (H&E) yang mempermudah dalam menganalisis keabnormalan struktur jaringan dan sel dengan memberikan kontras warna antar komponen sel (Wibowo & Maulani, 2024). Hasil dari analisis dan gambaran tersebut dapat mengidentifikasi struktur sel normal dan abnormal (Gawade et al., 2023).

Dengan kemajuan teknologi, citra histologi didigitalisasi menjadi *Whole Slide Imaging* (WSI), yang dimana WSI ini membantu sistem patologi digital dalam memberikan peluang untuk mengembangkan diagnosis awal penyakit dengan pendekatan algoritma CNN (Anisuzzaman et al., 2021). Berdasarkan beberapa penelitian sebelumnya telah digunakan beberapa jenis arsitektur dari algoritma CNN seperti *VGG19*, *InceptionV3*, dan *ResNet101* dalam diagnosis penyakit *Osteosarcoma* menggunakan citra histologi tulang *Osteosarcoma* (Anisuzzaman et al., 2021; Gawade et al., 2023). Meskipun model-model seperti *VGG19*, *InceptionV3*, dan *ResNet101* telah digunakan dalam diagnosis *Osteosarcoma*, model-model tersebut memiliki beberapa keterbatasan seperti kebutuhan komputasi yang tinggi, maka dibutuhkan beberapa eksplorasi untuk mengetahui model yang lebih efisien dan stabil dengan sumber daya komputasi yang terbatas. Maka dari itu, penelitian ini menggunakan model dari *deep learning*, yaitu model *EfficientNetV2* yang disarankan pada penelitian oleh Gawade et. Al (Gawade et al., 2023).

EfficientNetV2 merupakan salah satu arsitektur CNN yang dirilis pada tahun 2021 oleh *Google Brain*.

Model ini dilengkapi dengan operasi baru seperti *Fused-MBConv* untuk meningkatkan efisiensi dan kinerja model dalam melakukan *image processing* seperti deteksi dan klasifikasi gambar. Model ini memiliki beberapa jenis arsitektur, yaitu *EfficientNetV2-S*, *EfficientNetV2-M*, dan *EfficientNetV2-L*. *EfficientNetV2-S* merupakan varian terkecil yang dirancang untuk meningkatkan efisiensi dengan komputasi yang lebih kecil, *EfficientNetV2-M* memberikan keseimbangan antara dimensi dan kinerja tanpa memerlukan sumber daya komputasi yang besar, dan *EfficientNetV2-L* merupakan arsitektur model terbesar yang memaksimalkan akurasi dengan sumber daya komputasi yang besar (Tan & Le, 2021).

Meskipun model *EfficientNetV2* memberikan kinerja yang lebih baik, model ini memiliki beberapa kelemahan dalam kinerjanya. Maka, untuk membantu dalam peningkatan kinerja dan efisiensi model, perlu digunakan *optimizer*. *Optimizer* yang digunakan merupakan *Adam optimizer*. Meskipun *optimizer Adam* bukan yang terbaik, tetapi *optimizer* ini umumnya digunakan sebagai pilihan *default* untuk meningkatkan kinerja model (Vezakis et al., 2023).

Tujuan dari penelitian ini adalah untuk mengetahui dan menentukan kinerja akurasi dan stabilitas yang lebih unggul dengan membandingkan kinerja akurasi dan stabilitas dari ketiga model arsitektur *EfficientNetV2* dalam mendeteksi *Osteosarcoma* berdasarkan gambar histologi tulang dengan menggunakan *Adam Optimizer*.

2. METODE PENELITIAN

Metodologi penelitian yang dilakukan dimulai dengan pengumpulan dataset yang diperoleh dari website *The Cancer Imaging Archive (TCIA) Public Access* sebanyak 1144 gambar, lalu dataset tersebut akan dibagi menjadi beberapa *patch* gambar dengan ukuran 384x384 piksel yang sebelumnya berukuran 1024x1024 piksel. Lalu dari *patch* tersebut, akan dibagi menjadi 3 bagian yaitu data *train*, *validation*, dan *test* dengan perbandingan 80:10:10, 70:20:10, dan 60:30:10 untuk masing-masing data *train*, *validation*, dan *test*. Setelah melakukan *split* dataset, dilakukan augmentasi untuk data *patch train* dan *validation*. Setelah augmentasi dilakukan, data *train* dan *validation* akan dilatih (*training*) dengan menggunakan *base model* dari varian *EfficientNetV2*, ketiga model tersebut adalah *EfficientNetV2-S*, *EfficientNetV2-M*, *EfficientNetV2-L*, dan *optimizer Adam* digunakan untuk membantu meningkatkan kinerja model, *hyperparameter* yang digunakan berupa *learning rate*, *batch size*, *epoch*, *dropout*, dan *optimizer*. Setelah *training* dilakukan akan menghasilkan *best model* dari masing-masing model arsitektur yang digunakan, dan disimpan dalam *model H5*. Lalu, dari *best model H5* tersebut akan dilakukan *testing* atau pengujian dengan *input* gambar dari dataset *test*. Setelah dilakukan *testing* akan menghasilkan nilai metrik evaluasi dan melakukan

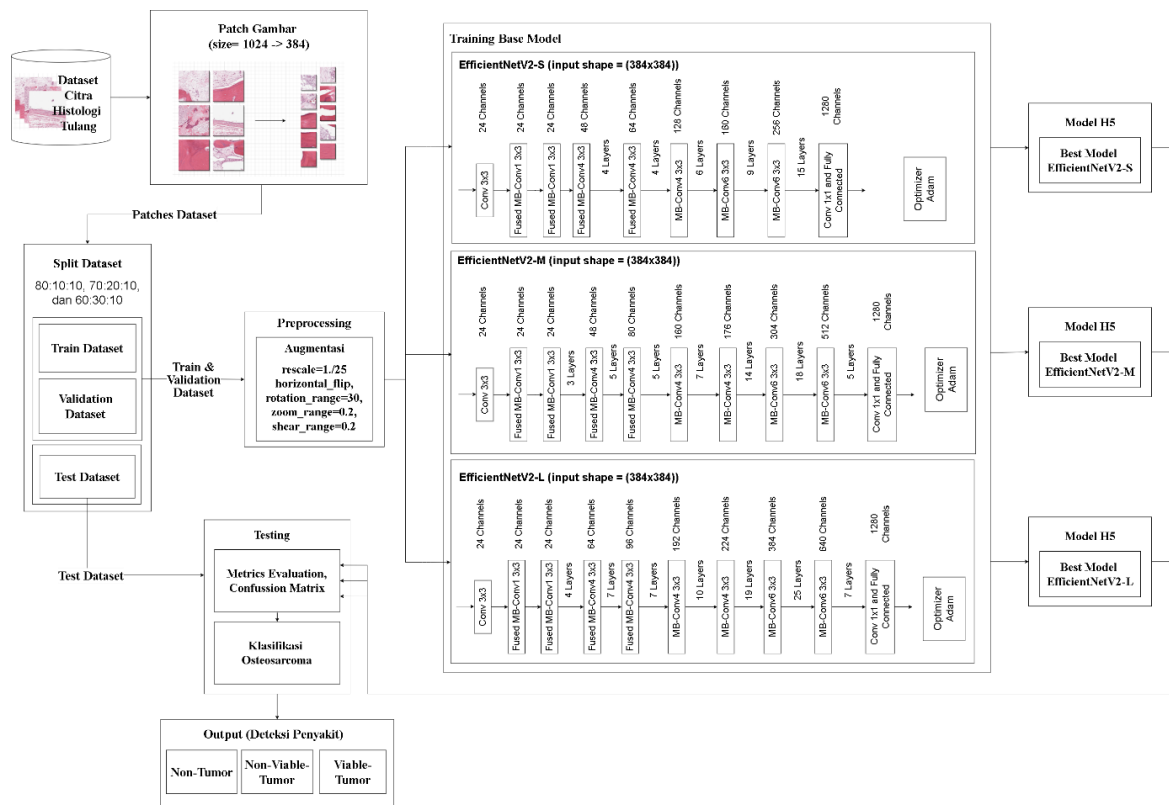
klasifikasi dari tiga *class*, yaitu *Non-Tumor*, *Non-Viable-Tumor*, dan *Viable-Tumor*. Setelah melakukan evaluasi kinerja berdasarkan metrik evaluasi, implementasi model dilakukan dengan aplikasi *GUI*. *Best* model terbaik digunakan untuk memprediksi gambar histologi yang diinputkan. Dalam penelitian ini, *input* citra baru diambil dari gambar asli dengan dimensi 1024x1024 piksel yang belum dipecah menjadi *patch* dan dari dataset uji.

Output yang dihasilkan adalah diagnosis awal citra histologi dari data *test* dengan dimensi 1024x1024 piksel dengan kelas yang sesuai, yaitu *Non-Tumor*, *Non-Viable-Tumor*, dan *Viable-Tumor* dan nilai probabilitas berupa angka dalam bentuk persen dari setiap model *EfficientNetV*. Berikut merupakan blok diagram alur dari penelitian ini yang ditunjukkan pada Gambar 1.

2.1. Pengumpulan Dataset

Data yang dikumpulkan merupakan dataset sekunder citra histologi tulang penyakit

Osteosarcoma dengan pewarnaan *Hematoxylin* dan *Eosin* (H&E) yang diperoleh dari website *The Cancer Imaging Archive (TCIA) Public Access* dengan link <https://www.cancerimagingarchive.net/collection/Osteosarcoma-tumor-assessment/>. Sampel arsip dikumpulkan oleh *clinical scientists of University of Texas Southwestern Medical Center, Dallas*. 4 pasien dari 50 pasien dipilih berdasarkan keragaman spesimen tumor setelah reseksi bedah. Pelabelan dataset dibagi menjadi tiga kelas, yaitu *Non-Tumor*, *Non-Viable-Tumor*, dan *Viable-Tumor*. Citra ini berukuran 1024x1024 piksel dengan resolusi yang telah diperbaiki (*Osteosarcoma-Tumor-Assessment | Osteosarcoma Data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment*, 2019). Berikut merupakan jumlah dan *sampel* dataset keseluruhan yang dapat dilihat pada Tabel 1 dan Tabel 2.



Gambar 1. Blok diagram proses diagnosis awal dan klasifikasi *Osteosarcoma*. Dataset diperoleh dari *The Cancer Imaging Archive (Osteosarcoma-Tumor-Assessment | Osteosarcoma Data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment*, 2019), dan arsitektur model berdasarkan literatur oleh (Sunil et al., 2022)

Tabel 1. Dataset(*Osteosarcoma-Tumor-Assessment | Osteosarcoma Data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment*, 2019)

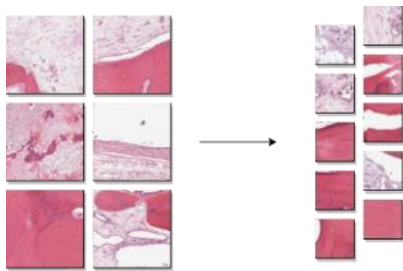
No	Kelas	Persen	Jumlah
1	<i>Non-Tumor</i>	47%	536
2	<i>Non-Viable-Tumor</i>	23%	263
3	<i>Viable-Tumor</i>	30%	345
Total			1144

Tabel 2. Sampel dataset(*Osteosarcoma-Tumor-Assessment | Osteosarcoma Data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment*, 2019)

NT	NVT	VT
		

2.2. Patch dataset

Patch dataset merupakan hasil dari pemotongan gambar asli, yang menghasilkan potongan-potongan kecil dari gambar. Tujuan membagi gambar jadi *patch* adalah untuk memberikan *input* gambar yang lebih kecil dengan ukuran yang disesuaikan dengan kebutuhan model agar model dapat menangkap informasi yang lebih terperinci dari setiap gambar pada dataset, sehingga dapat memudahkan model dalam mengidentifikasi dan mengklasifikasikan objek dan dapat meningkatkan kinerja dari model yang digunakan. (Arunachalam et al., 2019). Contoh pembagian gambar asli menjadi *patch* divisualisasikan pada Gambar 2. Lalu, setiap *patch* akan dianotasikan berdasarkan label.



Gambar 2. Patch dataset

2.3. Split dataset

Split dataset dilakukan pada dataset *patch* yang telah dianotasikan dengan label dan telah dikelompokkan sesuai dengan *class*. pembagian menjadi 3, yaitu *train*, *validation*, dan *test*. *Split* dataset yang dilakukan dari setiap *train*, *val*, dan *test* dibagi dengan perbandingan 80:10:10, 70:20:10, dan 60:30:10.

2.4. Augmentasi Data

Augmentasi data bertujuan untuk dilakukan untuk memperbanyak variasi data pelatihan dan generalisasi model (DLY et al., 2023). Proses atau jenis teknik augmentasi yang digunakan meliputi berupa *rescale*=1/255, *horizontal flip*, rotasi=30, *zoom*=0.2, dan *shear*=0.2.

2.5. Hyperparameter

Hyperparameter merupakan variable dikonfigurasi agar mendapatkan parameter terbaik (Susanto, 2023). Parameter yang dikonfigurasi dalam penelitian ini meliputi base dan jenis model, *batch size*, *learning rate*, *epoch*, *dropout*, *optimizer*, dan *Loss function*. *Hyperparameter* yang digunakan dapat dilihat pada Tabel 3.

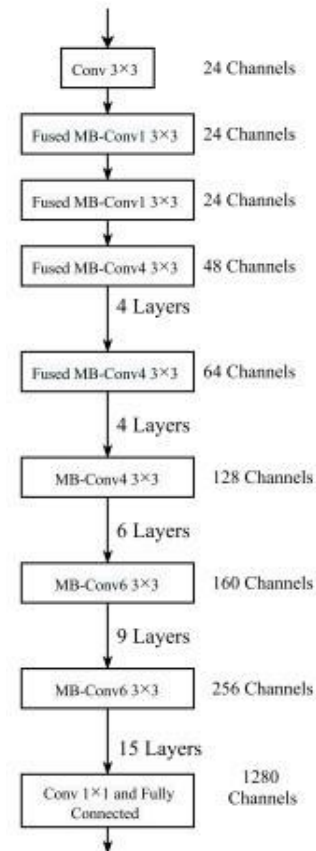
Tabel 3. Konfigurasi <i>hyperparameter</i>	
<i>Hyperparameter</i>	Konfigurasi
<i>Batch size</i>	32, 64
<i>Learning Rate</i>	0.001, 0.0001
<i>Epoch</i>	10, 20, 30, 40
<i>Dropout</i>	0.2

<i>Optimizer</i>	Adam
<i>Loss Function</i>	Categorical croosentropy

2.6. Model EfficientNetV2

EfficientNetV2 merupakan salah satu arsitektur CNN yang dirilis pada tahun 2021 oleh *Google Brain*. Model ini dilengkapi dengan operasi baru seperti *Fused-MBConv* untuk meningkatkan efisiensi dan kinerja model dalam melakukan *image processing* seperti deteksi dan klasifikasi gambar.

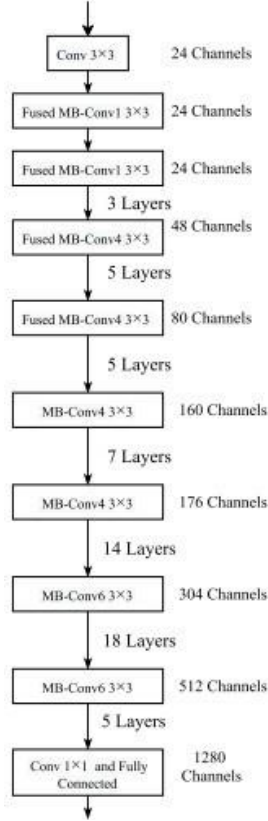
EfficientNetV2-S merupakan salah satu keluarga dari *EfficientNetV2* yang dirancang untuk melatih ukuran arsitektur yang ringkas, maka cocok untuk sumber daya yang terbatas. Arsitektur ini memprioritaskan efisiensi dengan menjaga keseimbangan kinerja, dimana dalam waktu yang bersamaan model ini memastikan untuk memberikan hasil yang optimal tetapi dengan ukuran yang lebih kecil dari arsitektur lainnya dan menekan fitur-fitur penting untuk mendapatkan hasil yang efisien dengan pengurangan ukuran hingga 6,8 kali. Arsitektur ini melibatkan penyesuaian regularisasi secara adaptif sehingga memungkinkan untuk mempercepat pelatihan tanpa mengorbankan akurasi (Tan & Le, 2021). Gambar 3. merupakan arsitektur dari *EfficientNetV2-S*.



Gambar 3. Arsitektur *EfficientNetV2-S* (Sunil et al., 2022)

EfficientNetV2-M merupakan arsitektur dengan ukuran yang lebih besar dari *EfficientNetV2-S* tetapi lebih kecil dari *EfficientNetV2-L*. Arsitektur ini memperoleh keseimbangan antara ukuran dan kinerja

tanpa memerlukan komputasi sebesar *EfficientNetV2-L*. Karena posisinya berada di tengah, arsitektur ini mencapai keseimbangan antara model dan kompleksitas yang membuat arsitektur ini serbaguna untuk berbagai aplikasi (Tan & Le, 2021). Gambar 4. merupakan arsitektur dari *EfficientNetV2-M*.

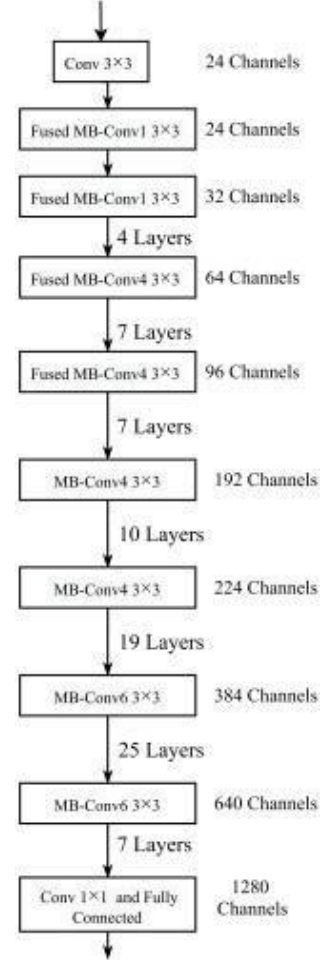


Gambar 4. Arsitektur *EfficientNetV2-M* (Sunil et al., 2022)

EfficientNetV2-L merupakan arsitektur terbesar dari ketiga varian *EfficientNetV2*, dimana arsitektur ini berfokus untuk memaksimalkan kinerja dan akurasi, memberikan kapasitas yang berpotensi lebih baik dari arsitektur lainnya, tapi waktu yang diperlukan lebih banyak dan sumber daya yang diperlukan lebih besar dan lebih kompleks selama proses pelatihan (Tan & Le, 2021). Gambar 5. merupakan arsitektur dari *EfficientNetV2-L*.

2.7. Adam Optimizer

Adaptive Moment Estimation (Adam) Optimizer, memiliki kemampuan secara otomatis untuk memperbarui bobot dan *learning rate*. Adaptabilitas tersebut dapat mengatasi masalah *sparse gradients* dan memanfaatkan rata-rata dari momen pertama dan momen kedua untuk mengadaptasi parameter *learning rate*.



Gambar 5. Arsitektur *EfficientNetV2-L* (Sunil et al., 2022)

Perhitungan *optimizer* ini dimulai dengan memanfaatkan momen pertama dan kedua, dimana setelah menghitung gradien, estimasi dari kedua momen akan mengalami bias yang diperbarui, persamaan tersebut dapat dirumuskan sebagai pada persamaan (1) dan (2).

$$m_t \leftarrow \beta_1 * m_{t-1} + (1 - \beta_1) * g_t \quad (1)$$

$$v_t \leftarrow \beta_2 * v_{t-1} + (1 - \beta_2) * g_t^2 \quad (2)$$

m_t dan v_t merupakan momen pertama dan kedua dari setiap waktu (t), β_1 dan β_2 adalah parameter dari laju momen pertama dan kedua, dan g_t merupakan gradien berdasarkan waktu (t). Setelah nilai diperbarui, momen terbaru digunakan untuk mengestimasi perbaruan parameter yang diterapkan, perumusan pada persamaan (3) dan (4).

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (3)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (4)$$

Dari persamaan (3) dan (4), dapat dilihat $\hat{m}_t$ dan $\hat{v}_t$ merupakan momen pertama dan momen kedua dari rumus yang memperbaiki bias awal momen pertama dan kedua, β_1 dan β_2 adalah parameter dari

laju momen pertama dan kedua berdasarkan waktu (t). Setelah itu, akan dilakukan perbaruan parameter model dengan rumus persamaan (5) dan (6)

$$\Delta\theta = -\eta \frac{m_t}{\sqrt{v_t + \epsilon}} * g_t \quad (5)$$

$$\theta_t \leftarrow \theta_{t-1} + \Delta\theta \quad (6)$$

(Syifa & Dewi, 2022).

2.8. Confussion Matrix

Confussion matrix merupakan tabel yang menunjukkan jumlah data uji yang diklasifikasikan dengan benar (*True Positive* dan *True negative*) dan jumlah data uji yang salah (*False Positive* dan *False Negative*) pengklasifikasiannya (Ezperanza et al., 2024). Gambar 6 menunjukkan tabel confussion matriks.

		Predicted	
		Positive	Negative
actual	Positive	TP	FP
	Negative	FN	TN

Gambar 6. Confussion matrix (Ezperanza et al., 2024)

Pada Gambar 6, terdapat 4 representasi proses klasifikasi *Confussion matrix*, dimana TP (*True Positive*) merupakan data positif dengan prediksi benar, TN (*True Negative*) adalah data *negative* dengan prediksi benar, FP (*False Positive*) adalah data *negative* prediksi positif, dan FN (*False negative*) yang adalah data positif prediksi data *negative* yang dirumuskan pada persamaan (7), (8), (9), dan (10).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$F1 = \frac{2*(precision*recall)}{(precision+recall)} \quad (10)$$

(Ezperanza et al., 2024).

3. HASIL DAN PEMBAHASAN

Tahap ini membahas mengenai performa dari masing-masing arsitektur model dalam mendeteksi penyakit *Osteosarcoma*.

3.1. Hasil Patch Dataset

Pembuatan *patch* dataset dihasilkan dari setiap gambar yang merupakan sampel lokal, diekstraksi dari gambar yang lebih besar menjadi lebih kecil. *Patch* yang dihasilkan dari ukuran gambar yang disesuaikan oleh kebutuhan model. Setiap *patch* yang dihasilkan akan dianotasikan dengan label berupa beberapa *file .csv* yang berisi nama *file* dari gambar

dan label dari nama *file* gambar tersebut. *Patch* yang dihasilkan sebanyak 18288 gambar untuk ukuran 384x384 piksel.

3.2. Hasil Split Dataset

Split dataset dilakukan dengan pembagian rasio *split* 80:10:10, 70:20:10, dan 60:30:10. Tabel 4. memperlihatkan hasil dari setiap rasio *split* dataset.

Tabel 4. Hasil *split* dataset 80:10:10

Dimensi Patch	Rasio Split	Jumlah Gambar
384x384	80%	14624
	10%	1824
	10%	1840
Total		18288

Tabel 5. Hasil *split* dataset 70:20:10

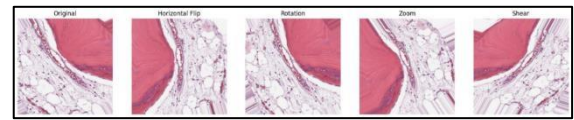
Dimensi Patch	Rasio Split	Jumlah Gambar
384x384	70%	12800
	20%	3648
	10%	1840
Total		18288

Tabel 6. Hasil *split* dataset 60:30:10

Dimensi Patch	Rasio Split	Jumlah Gambar
384x384	60%	10960
	30%	5488
	10%	1840
Total		18288

3.3. Hasil Augmentasi Data

Proses ini dapat memperkaya dataset sebelum dilatih sehingga Model dapat mempelajari data yang beragam. Proses atau jenis teknik augmentasi yang digunakan meliputi berupa *rescale=1/255*, *horizontal flip*, rotasi=30, *zoom=0.2*, dan *shear=0.2*. Gambar 7. memperlihatkan hasil dari augmentasi data.



Gambar 7. Hasil augmentasi data

3.4. Skenario Training EfficientNetV2

Penelitian ini menggunakan model *EfficientNetV2* untuk melakukan pelatihan (*training*) klasifikasi dan deteksi pada penyakit *Osteosarcoma*. Penerapan Model *EfficientNetV2* dengan hasil *hyperparameter* terbaik pada *batch* 32 dan 64, *learning rate* 0.001 dan 0.0001, *dropout* 0.2, dan *optimizer Adam*, dimana hasil *best* model dari masing-masing model dengan konfigurasi *hyperparameter* yang diterapkan dengan E adalah *epoch*, B adalah *batch*, LR adalah *learning rate*, TA adalah *Train Accuracy*, TL adalah *Train Loss*, VA adalah *Val Accuracy*, dan VL adalah *Val Loss* berdasarkan *trade off* terbaik dari masing-masing *learning rate*. Tabel 7, Tabel 8, dan Tabel 9 menunjukkan hasil *Best Model* dari Model *EfficientNetv2-S*, *EfficientV2-M*, dan *EfficientNetV2-L*.

Tabel 7. Best Model *EfficientNetV2-S*

Konfigurasi Hyperparameter terbaik							
Split Data	E	B	LR	TA	TL	VA	VL
80:10:10	30	32	0.0001	0.8887	0.2948	0.8766	0.3491
	30	64	0.001	0.8988	0.2723	0.8821	0.3205
70:20:10	30	32	0.0001	0.8920	0.2866	0.8857	0.3099
	40	32	0.001	0.8998	0.2633	0.8991	0.2803
60:30:10	30	32	0.0001	0.8979	0.2811	0.8870	0.3231
	30	64	0.001	0.9036	0.2552	0.8936	0.3015

Tabel 8. Best Model *EfficientNetV2-M*

Konfigurasi Hyperparameter terbaik							
Split Data	E	B	LR	TA	TL	VA	VL
80:10:10	40	32	0.0001	0.8838	0.3171	0.8745	0.3553
	40	64	0.001	0.8826	0.3181	0.8821	0.3362
70:20:10	40	32	0.0001	0.8813	0.3183	0.8879	0.3248
	20	64	0.001	0.8891	0.2985	0.8936	0.3068
60:30:10	40	32	0.0001	0.8828	0.3151	0.8761	0.3588
	30	64	0.001	0.8896	0.2870	0.8863	0.3277

Tabel 9. Best Model *EfficientNetV2-L*

Konfigurasi Hyperparameter terbaik							
Split Data	E	B	LR	TA	TL	VA	VL
80:10:10	40	32	0.0001	0.8723	0.3471	0.8728	0.3787
	40	64	0.001	0.8877	0.3087	0.8942	0.3417
70:20:10	40	32	0.0001	0.8745	0.3369	0.8914	0.3245
	40	32	0.001	0.8863	0.3040	0.9013	0.2980
60:30:10	40	32	0.0001	0.8825	0.3274	0.8816	0.3555
	20	32	0.001	0.8865	0.3050	0.8923	0.3301

3.5. Evaluasi Kinerja

Evaluasi kinerja model dilakukan dengan mengukur akurasi dan stabilitas berdasarkan masing-masing best model varian *EfficientNetV2* dengan *learning rate* yang berbeda berdasarkan rasio *split* dataset yang digunakan. Akurasi didapatkan berdasarkan jumlah prediksi yang benar dibandingkan dengan total data uji. Stabilitas model diperoleh dari nilai *F1-Score* yang merupakan gabungan dari nilai *precision* dan *recall* untuk setiap kelas. Model dengan nilai *F1-Score* yang stabil antar kelas dianggap memiliki kinerja yang lebih konsisten. Evaluasi ini tidak hanya membandingkan akurasi tertinggi, tetapi mempertimbangkan juga mengenai *trade-off* antar stabilitas model yang menjadi faktor utama untuk pemilihan model terbaik dalam diagnosis awal dan klasifikasi *Osteosarcoma*.

Tabel 10, Tabel 11, dan Tabel 12 merupakan hasil evaluasi kinerja dari masing-masing model *EfficientNetV2* dengan Split merupakan Split Dataset, LR (*learning rate*), E (*Epoch*), BS (*Batch Size*), C (*Class*), P (*Precision*), R (*Recall*), F1 (*F1-Score*), dan Acc (Akurasi).

Tabel 10. Evaluasi Kinerja *EfficientNetV2-S*

Split	LR	C	P	R	F1	Acc
80:10:10	0.0001	NT	0.86	0.94	0.90	88.42%
		NVT	0.91	0.85	0.88	
		VT	0.92	0.81	0.86	
	0.001	NT	0.88	0.92	0.90	88.86%
		NVT	0.91	0.86	0.88	
		VT	0.89	0.85	0.87	
70:20:10	0.0001	NT	0.83	0.92	0.87	86.47%
		NVT	0.89	0.80	0.84	
10	0.0001	NT	0.86	0.94	0.90	88.42%
		NVT	0.91	0.85	0.88	

Split	LR	C	P	R	F1	Acc
80:10:10	0.001	NT	0.85	0.91	0.88	87.39%
		NVT	0.89	0.81	0.85	
		VT	0.89	0.87	0.88	
	0.0001	NT	0.84	0.92	0.88	86.58%
		NVT	0.83	0.88	0.85	
		VT	0.94	0.79	0.86	
60:30:10	0.001	NT	0.86	0.92	0.88	87.01%
		NVT	0.82	0.90	0.85	
10	0.001	NT	0.86	0.92	0.88	87.01%
		NVT	0.82	0.90	0.85	

Tabel 11. Evaluasi Kinerja *EfficientNetV2-M*

Split	LR	C	P	R	F1	Acc
80:10:10	0.0001	NT	0.87	0.94	0.90	88.80%
		NVT	0.94	0.86	0.90	
		VT	0.87	0.83	0.85	
	0.001	NT	0.86	0.94	0.90	87.66%
		NVT	0.94	0.80	0.87	
		VT	0.84	0.84	0.84	
70:20:10	0.0001	NT	0.83	0.92	0.87	86.03%
		NVT	0.89	0.78	0.83	
		VT	0.89	0.83	0.86	
	0.001	NT	0.84	0.92	0.87	86.03%
		NVT	0.89	0.78	0.83	
		VT	0.89	0.83	0.86	
60:30:10	0.0001	NT	0.84	0.93	0.88	86.03%
		NVT	0.83	0.86	0.85	
		VT	0.92	0.77	0.84	
	0.001	NT	0.86	0.93	0.89	87.50%
		NVT	0.84	0.87	0.86	
		VT	0.93	0.81	0.87	

Tabel 12. Evaluasi Kinerja *EfficientNetV2-L*

Split	LR	C	P	R	F1	Acc
80:10:10	0.0001	NT	0.87	0.92	0.90	87.50%
		NVT	0.89	0.83	0.86	
		VT	0.87	0.83	0.85	
	0.001	NT	0.80	0.90	0.89	88.15%
		NVT	0.88	0.86	0.87	
		VT	0.87	0.87	0.87	
70:20:10	0.0001	NT	0.83	0.91	0.87	85.60%
		NVT	0.87	0.78	0.83	
		VT	0.89	0.82	0.85	
	0.001	NT	0.85	0.89	0.87	86.20%
		NVT	0.85	0.82	0.84	
		VT	0.89	0.84	0.87	
60:30:10	0.0001	NT	0.86	0.93	0.89	85.87%
		NVT	0.80	0.85	0.82	
		VT	0.92	0.77	0.84	
	0.001	NT	0.87	0.91	0.89	86.63%
		NVT	0.79	0.88	0.83	
		VT	0.93	0.80	0.86	

Berdasarkan Tabel 10 - 12 dengan *class* NT (*Non-Tumor*), NVT (*Non-Viable-Tumor*), dan VT (*Viable-Tumor*), analisis hasil evaluasi kinerja secara keseluruhan dapat dilihat pada Tabel 13.

Tabel 13. Evaluasi Kinerja Model Terbaik

Model	SD	LR	E	BS	C	P	R	F1	Acc
EffNet V2-S	80:10:10	0.001	30	64	NT	0.88	0.92	0.90	88.86%
					NVT	0.91	0.86	0.88	
					VT	0.89	0.85	0.87	
EffNet V2-M	80:10:10	0.0001	40	32	NT	0.87	0.94	0.90	88.80%
					NVT	0.94	0.86	0.90	
					VT	0.87	0.83	0.85	

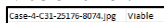
Model	SD	LR	E	BS	C	P	R	F1	Acc
EffNet 80:10:10 V2-L	0.001	40	64		NT	0.80	0.90	0.89	88.15%
					NVT	0.88	0.86	0.87	
					VT	0.87	0.87	0.87	

Jika dilihat pada Tabel 13, model terbaik yang dipilih didasari oleh dua kriteria yaitu akurasi dan stabilitas. Model *EfficientNetV2-S*, *split* dataset 80:10:10, dengan *learning rate* 0.001, *epoch* 30, *batch size* 64 dipilih menjadi model terbaik karena mendapatkan hasil akurasi tertinggi yaitu 88.86%. Tetapi, jika dilihat berdasarkan stabilitas model, model *EfficientNetV2-M* pada *split* dataset 80:10:10, dengan *learning rate* 0.0001, *epoch* 40, dan *batch dimensi* 32, menghasilkan nilai *F1-Score* yang lebih stabil dibandingkan dengan Model *EfficientNetV2-S*, *split* dataset 80:10:10, dengan *learning rate* 0.001, *epoch* 30, *batch size* 64. Akurasi dari Model *EfficientNetV2-M* pada *split* dataset 80:10:10, dengan *learning rate* 0.0001, *epoch* 40, dan *batch size* 32 juga menghasilkan akurasi yang tinggi yaitu sebesar 88.80%.

4. Pengujian Model

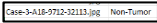
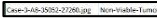
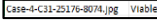
Pengujian model dilakukan dengan menggunakan hasil dari nilai evaluasi kinerja terbaik yang diperoleh dari masing-masing model *EfficientnetV2*. Tabel 14 menunjukkan hasil pengujian sistem dengan menggunakan *Best Model EfficientNetV2-S*.

Tabel 14. Pengujian *EfficientNetV2-S*

No	Skenario Pengujian	Hasil Prediksi	Validasi
1.	Input Kelas Non-Tumor	Hasil yang diharapkan adalah Non-Tumor.  Case-3-A7-10464-40567.jpg Non-Tumor with 90.50% confidence	Hasil klasifikasi benar menunjukkan bahwa gambar citra histologi yang di <i>inputkan</i> adalah kelas Non-Tumor.  Case-3-A18-0712-32113.jpg Non-Tumor
2.	Input Kelas Non-Viable-Tumor	Hasil yang diharapkan adalah Non-Viable-Tumor.  Case-3-A8-35052-27260.jpg Non-Viable Tumor with 80.41% confidence	Hasil klasifikasi benar menunjukkan bahwa gambar citra histologi yang di <i>inputkan</i> adalah kelas Non-Viable-Tumor.  Case-3-A8-35052-27260.jpg Non-Viable-Tumor
3.	Input Kelas Viable-Tumor	Hasil yang diharapkan adalah Viable-Tumor.  Case-4-C31-2517-6-8074.jpg Viable-Tumor with 93.51% confidence	Hasil klasifikasi benar menunjukkan bahwa gambar citra histologi yang di <i>inputkan</i> adalah kelas Viable-Tumor.  Case-4-C31-2517-6-8074.jpg Viable

Tabel 15 menunjukkan hasil pengujian sistem dengan menggunakan *Best Model EfficientNetV2-M*.

Tabel 15. Pengujian *EfficientNetV2-M*

No	Skenario Pengujian	Hasil Prediksi	Validasi
1.	Input Kelas Non-Tumor	Hasil yang diharapkan adalah Non-Tumor.  Case-3-A7-10464-40567.jpg Non-Tumor with 88.81% confidence	Hasil klasifikasi benar menunjukkan bahwa gambar citra histologi yang di <i>inputkan</i> adalah kelas Non-Tumor.  Case-3-A18-0712-32113.jpg Non-Tumor
2.	Input Kelas Non-Viable-Tumor	Hasil yang diharapkan adalah Non-Viable-Tumor.  Case-3-A8-35052-27260.jpg Non-Viable Tumor with 73.04% confidence	Hasil klasifikasi benar menunjukkan bahwa gambar citra histologi yang di <i>inputkan</i> adalah kelas Non-Viable-Tumor.  Case-3-A8-35052-27260.jpg Non-Viable-Tumor
3.	Input Kelas Viable-Tumor	Hasil yang diharapkan adalah Viable-Tumor.  Case-4-C31-2517-6-8074.jpg Viable Tumor with 89.08% confidence	Hasil klasifikasi benar menunjukkan bahwa gambar citra histologi yang di <i>inputkan</i> adalah kelas Viable-Tumor.  Case-4-C31-2517-6-8074.jpg Viable

Dapat dilihat pada tabel 14 dan tabel 15, Model *EfficientNetV2-S* dan *EfficientNetV2-M* mendapatkan hasil prediksi dari setiap kelas dengan *validasi* pada *file csv pathologistvalidation* dengan benar.

5. KESIMPULAN

Dalam penelitian ini, hasil yang didapatkan untuk kriteria nilai akurasi terbaik diraih oleh *EfficientNetV2-S* dengan *split* dataset 80:10:10, *epoch* 30, *learning rate* 0.001, dan *batch size* 64 mendapatkan akurasi sebesar 88.86%. Tetapi jika dilihat dari kinerja stabilitas, Model *EfficientNetV2-M* dengan *split* dataset 80:10:10, *learning rate* 0.0001, *epoch* 40, dan *batch size* 32, memberikan kinerja yang lebih satbil dengan nilai *F1-Score* yang lebih baik dan menghasilkan akurasi yang cukup tinggi sebesar 88.80%.

Secara keseluruhan, model *EfficientNetV2-S* dengan *split* dataset 80:10:10, *epoch* 30, *learning rate* 0.001, dan *batch size* 64, dan akurasi sebesar 88.86% dianggap menjadi model yang optimal pada penelitian ini, model ini mendapatkan hasil akurasi tertinggi dan efisiensi yang lebih baik berdasarkan ukuran model.

Untuk penelitian selanjutnya, disarankan untuk memperbanyak jumlah dan variasi dataset dengan penambahan jumlah citra histologi dan variasi teknik augmentasi untuk meningkatkan kemampuan generalisasi model, lalu melakukan penambahan variasi ukuran dimensi gambar yang disesuaikan dengan kebutuhan model untuk meningkatkan hasil kinerja dari setiap model yang digunakan.

DAFTAR PUSTAKA

- ANHAR, A., & PUTRA, R. A. 2023. Perancangan dan Implementasi Self-Checkout System pada Toko Ritel menggunakan Convolutional Neural Network (CNN). *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 11(2), 466. <https://doi.org/10.26760/elkomika.v11i2.466>
- ANISUZZAMAN, D. M., BARZEKAR, H., TONG, L., LUO, J., & YU, Z. 2021. A deep learning study on osteosarcoma detection from histological images. *Biomedical Signal Processing and Control*, 69(July), 102931. <https://doi.org/10.1016/j.bspc.2021.102931>
- ARUNACHALAM, H. B., MISHRA, R., DAESCU, O., CEDERBERG, K., RAKHEJA, D., SENGUPTA, A., LEONARD, D., HALLAC, R., & LEAVEY, P. 2019. Viable and necrotic tumor assessment from whole slide images of osteosarcoma using machine-learning and deep-learning models. *PLoS ONE*, 14(4), 1–19. <https://doi.org/10.1371/journal.pone.0210706>
- DLY, I. A., JASRIL, J., SANJAYA, S., HANDAYANI, L., & YANTO, F. 2023. Klasifikasi Citra Daging Sapi dan Babi Menggunakan CNN Alexnet dan Augmentasi Data. *Journal of Information System Research (JOSH)*, 4(4), 1176–1185. <https://doi.org/10.47065/josh.v4i4.3702>
- EZPERANZA, M., LETIK, F., & BISILISIN, F. Y. 2024. Klasifikasi Tanaman Herbal Berdasarkan Tekstur Daun Menggunakan Backpropagation Berbasis Citra. 01(06), 1–10. <https://doi.org/https://doi.org/10.70404/ketik.v1i06.87>
- GAWADE, S., BHANSALI, A., PATIL, K., & SHAIKH, D. 2023. Application of the convolutional neural networks and supervised deep-learning methods for osteosarcoma bone cancer detection. *Healthcare Analytics*, 3(February), 100153. <https://doi.org/10.1016/j.health.2023.100153>
- GUNAWAN, D., & SETIAWAN, H. 2022. Convolutional Neural Network dalam Citra Medis. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 2(2), 376–390. <https://doi.org/10.24002/konstelasi.v2i2.5367>
- MAWARDI, D. P., Novita, M., & Dwi Saputro, N. 2024. Deteksi Awal Klasifikasi Jenis Penyakit Kanker Kulit Dengan Algoritma Convolutional Neural Network (CNN) Berbasis Mobile Apps. 3(2), 1–6. <https://doi.org/https://doi.org/10.30872/atasi.v3i2.2305>
- Osteosarcoma-Tumor-Assessment | Osteosarcoma data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment*. (2019). <https://www.cancerimagingarchive.net/collecton/osteosarcoma-tumor-assessment/>
- SUNIL, C. K., JAIDHAR, C. D., & PATIL, N. 2022. Cardamom Plant Disease Detection Approach Using EfficientNetV2. *IEEE Access*, 10, 789–804. <https://doi.org/10.1109/ACCESS.2021.3138920>
- SUSANTO, L. A. 2023. Pemilihan Hyperparameter Pada Alexnet Cnn Untuk Klasifikasi Citra Penyakit Kedelai. *Indexia*, 5(02), 113. <https://doi.org/10.30587/indexia.v5i02.5508>
- SYIFA, S. A., & DEWI, I. A. 2022. Arsitektur Resnet-152 dengan Perbandingan Optimizer Adam dan RMSProp untuk Mendeteksi Penyakit Paru-Paru. *Journal MIND Journal | ISSN*, 7(2), 139–150. <https://doi.org/10.26760/mindjournal.v7i2.139-150>
- TAN, M., & LE, Q. V. 2021. EfficientNetV2: Smaller Models and Faster Training. *Proceedings of Machine Learning Research*, 139, 10096–10106.
- VEZAKIS, I. A., LAMBROU, G. I., & MATSOPOULOS, G. K. 2023. Deep Learning Approaches to Osteosarcoma Diagnosis and Classification: A Comparative Methodological Approach. *Cancers*, 15(8), 1–15. <https://doi.org/10.3390/cancers15082290>
- WIBOWO, T., & MAULANI, Y. 2024. Perbedaan Hasil Pewarnaan Hematoxylin-Eosin Preparat Limfonodi Pada Proses Clearing Menggunakan Xylol Dan Minyak Zaitun. 1(3), 197–201. <https://doi.org/https://doi.org/10.37985/plenar.yhealth.v1i3.562>

Halaman ini sengaja dikosongkan