
PERBANDINGAN OPTIMASI SGD, ADADELTA, DAN ADAM DALAM KLASIFIKASI HYDRANGEA MENGUNAKAN CNN

Desi Irfan¹, Rika Rosnelly², Masri Wahyuni³, Jaka Tirta Samudra⁴,
Aditia Ranga⁵

Universitas Potensi Utama, Medan

e-mail: ¹desiirfan@gmail.com, ²rikarosnelly@gmail.com, ³masriwahyuni997@gmail.com,
⁴jakatirta135@gmail.com, ⁵aditiaranga567@gmail.com

Abstract: *invasive species are threatening indigenous species habitat in many countries around the world. Nowadays, the monitoring method relies on scientists. Scientists are skilled to see the determined areas and record the living species. Applying high skill labors requires high cost, inefficient time and limited scope as the large area cannot be reached by the man. In this research, engine based learning approach was presented to identify the image of invasive hydrangea (indigenous species from Asia) with data collection around 800 images taken from the Brazil national forest and Hydrangea appears in some images. Gradient Descent optimization method is frequently used for artificial neural network. This method roles to discover standard grade for the best output. The Gradient Descent method role play is minimizing the cost function grade by changing the parameter grade step by step. Three optimization methods have been implemented namely Stochastic Gradient Descent (SGD), ADADELTA, and Adam in the artificial neural network (Ann) for classifying aritmia data [32]. This research used the most suitable error grade limitation from each optimization method as the indicators at the end of the training. The result of this research showed that artificial nerve tissue using Adam optimization gets the highest accuracy compared with SDG and ADADELTA optimization methods. Deep Learning Technique applied extensively in image introduction is Adam optimization. The training model has reached accuracy to 83, 5 % and showed properness of approach conducted.*

Keywords: SGD, Adadelata, Adam, Optimizer Function

Abstrak: Spesies invasif mengancam habitat spesies asli di banyak negara di dunia. Saat ini dalam metode pemantauan mereka tergantung pada pengetahuan ahli. Ilmuwan terlatih mengunjungi area yang ditentukan dan mencatat spesies yang menghuninya. Menggunakan tenaga kerja berkualifikasi tinggi seperti itu membutuhkan biaya yang mahal, tidak efisien waktu dan jangkauan yang terbatas karena manusia tidak dapat mencakup area yang luas. Dalam makalah ini, pendekatan berbasis pembelajaran mesin disajikan untuk mengidentifikasi gambar hydrangea invasif (spesies invasif asli Asia) dengan kumpulan data yang berisi sekitar 800 gambar yang diambil di hutan nasional Brasil dan di beberapa gambar terdapat Hydrangea. Metode optimasi Gradient Descent sering digunakan untuk pelatihan Jaringan Syaraf Tiruan (JST). Metode ini berperan dalam menemukan nilai bobot yang memberikan nilai keluaran terbaik. Prinsip kerja metode Gradient Descent adalah memperkecil nilai fungsi biaya dengan mengubah nilai parameter selangkah demi selangkah. Telah diimplementasikan tiga buah metode optimasi yaitu Stochastic Gradient Descent (SGD), ADADELTA, dan Adam pada sistem Jaringan Saraf Tiruan untuk klasifikasi data aritmia [32]. Penelitian ini menggunakan batas nilai kesalahan yang paling sesuai dari masing-masing metode optimasi sebagai kriteria pemberhentian pelatihan. Hasil penelitian menunjukkan Jaringan Saraf Tiruan dengan optimasi Adam menghasilkan akurasi tertinggi dibandingkan dengan dengan metode optimasi SGD dan ADADELTA. Teknik Deep Learning yang diterapkan secara ekstensif pada pengenalan gambar yang digunakan memanfaatkan metode optimizer

Adam . Model yang kita latih menggunakan fungsi optimisasi Adam mencapai akurasi 83,5% pada tes yang dilakukan, menunjukkan kelayakan pada pendekatan yang dilakukan

Kata kunci: SGD, Adadelta, Adam, Fungsi Optimisasi

PENDAHULUAN

Di Indonesia, Serangan organisme pengganggu tanaman (OPT) menjadi salah satu permasalahan yang sering dialami petani. Sementara alang-alang sekarang menyebar ke seluruh bagian Tenggara, menggantikan tanaman asli. Menurut Rahayu (1993) dalam Sianturi (2003), luas padang alang-alang di Indonesia diperkirakan sebesar 10.56 juta H.. Ini tidak memberikan nilai makanan bagi satwa liar asli, dan meningkatkan ancaman kebakaran hutan karena alang-alang lebih cepat tumbuh daripada rumput asli. Ini baru satu spesies invasif dari banyak spesies yang dapat memiliki konsekuensi merusak pada lingkungan, ekonomi, dan bahkan kesehatan manusia. Meskipun berdampak luas, upaya untuk melacak lokasi dan penyebaran spesies invasif terus dilakukan tetapi biaya yang sangat mahal mengakibatkan sulitnya melakukan dalam skala besar. Karena para ilmuwan tidak dapat mengambil sampel area dalam jumlah besar, beberapa teknik pembelajaran mesin digunakan untuk memprediksi ada tidaknya spesies invasif di area yang belum diambil sampelnya. Keakuratan pendekatan ini jauh dari optimal, tetapi masih memberikan kontribusi pada pendekatan untuk memecahkan masalah ekologi. Para ahli ekologi telah menyelenggarakan kompetisi Kaggle untuk memeriksa apakah komputer dapat secara akurat mengidentifikasi spesies invasif dalam gambar. Dalam karya ini, kami menunjukkan bahwa pembelajaran transfer Deep Convolutional Neural Network (CNNs Trained on imagenet dataset) bekerja dengan baik dalam mengidentifikasi gambar hydrangea invasive hal ini juga tidak terlepas pada fungsi optimizer yang digunakan. Dalam

computer vision, CNN telah dikenal sebagai model visual yang kuat serta menghasilkan hierarki fitur segmentasi yang akurat. Mereka juga dikenal melakukan prediksi yang relatif lebih cepat daripada algoritma lain dengan tetap mempertahankan kinerja kompetitif pada waktu yang sama. Metode optimisasi Gradient Descent sering digunakan untuk pelatihan Jaringan Syaraf Tiruan (JST). Metode ini berperan dalam menemukan nilai bobot yang memberikan nilai keluaran terbaik. Prinsip kerja metode Gradient Descent adalah memperkecil nilai fungsi biaya dengan mengubah nilai parameter selangkah demi selangkah. Telah diimplementasikan tiga buah metode optimisasi yaitu SGD, ADADELTA, dan Adam pada sistem Jaringan Saraf Tiruan untuk klasifikasi data aritmia. Pada penelitian ini metode Adam terbukti lebih akurat dibandingkan kedua metode lainnya diatas. Ini sesuai dengan penelitian sebelumnya tentang “Perbandingan Metode Optimisasi Stochastic Gradient Descent, ADADELTA, dan Adam pada Jaringan Saraf Tiruan dalam Klasifikasi Data Aritmia”.

Spesies invasif adalah spesies yang bukan habitat asli wilayah tertentu (spesies yang diperkenalkan), dan yang memiliki kecenderungan untuk menyebar ke tingkat yang diyakini menyebabkan kerusakan pada lingkungan, ekonomi manusia atau kesehatan manusia . Kriteria untuk spesies invasif ini menjadi kontroversial, karena persepsi yang sangat berbeda terjadi di antara para peneliti yang menimbulkan keprihatinan dengan subjektivitas istilah "invasif". Sementara semua spesies bersaing untuk bertahan hidup, spesies invasif tampaknya memiliki sifat-sifat tertentu atau kombinasi tertentu dari sifat-sifat yang memungkinkan mereka untuk bersaing

dengan spesies asli. Dalam beberapa kasus, persaingan tingkat pertumbuhan dan reproduksi. Dalam kasus lain, spesies berinteraksi satu sama lain lebih secara berkesinambungan. Para peneliti tidak setuju dengan ciri-ciri sebagai penanda tumbuhan invasif. Satu studi menemukan bahwa dari daftar spesies invasif dan noninvasif, 86% dari spesies invasif dapat diidentifikasi dari ciri-ciri saja di Lain penelitian menemukan spesies invasif cenderung hanya memiliki sedikit bagian dari sifat-sifat yang dapat diduga dan banyak sifat serupa ditemukan pada spesies noninvasif, disini kita membutuhkan penjelasan.

Ciri-ciri spesies invasif yang umum adalah sebagai berikut:

1. Pertumbuhan cepat.
2. Reproduksi cepat.
3. Kemampuan penyebaran yang tinggi.
4. Plastisitas fenotipe (kemampuan untuk mengubah pertumbuhan) bentuk yang sesuai dengan kondisi saat ini).
5. Adaptasi terhadap kondisi di berbagai lingkungan (kompetensi ekologi).
6. Kemampuan bertahan hidup dari berbagai jenis makanan (generalis).
7. Association with humans (Terbiasa dilihat manusia).
8. Prior successful invasions (keberhasilan menguasai suatu area).

METODE

Pada bagian ini, kami menggunakan prosedur penelitian pada gambar 1 berikut ini



Data Preprocessing Teknik pra-pemrosesan digunakan untuk mempersiapkan citra agar dapat menghasilkan ciri yang lebih baik pada tahap pemisahan ciri terhadap proses pengenalan pola.

Kumpulan data yang disediakan oleh Kaggle memiliki 2295 data training dan 1531 data tes. Semua gambar berukuran dalam ukuran 866 × 1154 piksel. Sampel positif yang khas dapat dilihat pada *Gambar. 2*. Sekitar dua pertiga dari data training adalah positif (dari sebaran sampel dapat dilihat pada *Gambar.3*). Ada sejumlah besar gambar yang hanya memiliki porsi kecil dari gambar yang ditempati oleh *hydrangea*. Gambar negatif sebagian besar adalah pemandangan hutan tanpa *hydrangea*, seperti yang dapat dilihat pada *gambar4*. Kadang-kadang mereka mengandung fitur lain seperti danau, kuda, gudang, atau orang.



Gambar Sampel Negative



Gambar Sampel Positif

Augmentasi data adalah suatu proses dalam pengolahan data gambar,

augmentasi merupakan proses mengubah atau memodifikasi gambar sedemikian rupa sehingga komputer akan mendeteksi bahwa gambar yang diubah adalah gambar yang berbeda, namun manusia masih dapat mengetahui bahwa gambar yang diubah tersebut adalah gambar yang sama.

Untuk memaksimalkan beberapa contoh pelatihan kami dan meningkatkan akurasi model, kami menambah data melalui sejumlah transformasi acak. yang dipilih teknik augmentasi data adalah:

1. *Transform random rotation* (15°), bertujuan untuk memperkaya metode pelatihan, karena setiap data ditarik walaupun sejenis akan memberikan posisi gambar yang berbeda.
2. *resize crop scale* = (80%), dilakukan untuk menyeragamkan ukuran daripada gambar yang berbeda-beda.
3. *Horizontal flip* acak, sama halnya dengan *transform random rotation* hanya pada bagian ini berbebtuk *flip*.
4. Transformasi Tensor, agar dapat dibaca saat pembuatan model.



Gambar rotation acak (15°)



Gamba Crop 80%



Gambar Flip (gambar kiri kekanan)

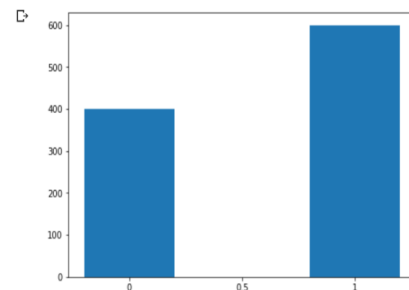
```
3 tensor.shape
4
↳ torch.Size([128, 3, 64, 64])
```

Gambar (Tensor)

Dari *gambar*.diatas dapat dilihat bahwa tensor memaparkan format N,C,H,W yaitu Bitch Size 128 gambar dengan 3 Chanel dan ukuran image 64x64. Selanjutnya diharapkan bahwa augmentasi data juga harus menghindari overfitting, serta kemampuan untuk meningkatkan dalam generalisasi model.

Data Splitting Salah satu keputusan pertama yang harus diambil ketika memulai proyek pemodelan adalah bagaimana memanfaatkan data yang ada. Salah satu teknik umum adalah membagi data menjadi dua kelompok yang biasanya disebut sebagai set pelatihan dan pengujian.

Pada penelitian ini kami menggunakan 1000 data yang terdiri dari 600 data training dan 400 data tes.



Distribusi Probabilitas

Arsitektur dan Konfigurasi
Menurut Francis DK Ching (1979), Arsitektur membentuk suatu tautan yang mempersatukan ruang, bentuk, teknik dan fungsi[18]. Pada tahap ini kita akan membuat arsitektur yang terdiri dari 4 lapisan convolutional layer, Relu layer, padding, dan max pooling pada *feature learning*. Sedangkan pada lapisan classification terdapat 1 lapisan flatten, dan 3 lapisan neural network atau fully connected yaitu (*Linear*, Relu layer dan Dropout 10%). Pada lapisan softmax terdiri output.

Setelah arsitektur model CNN dibuat, maka tahap selanjutnya adalah config. Config berfungsi untuk mengatur semua parameter yang akan digunakan pada model yang akan dibangun. Beberapa *hyperparameter* yang akan digunakan untuk mendapatkan model dengan tingkat akurasi yang terbaik dalam model CNN tersebut dapat dilihat pada Tabel 1.

Tabel 1 Config Hyperparameter Arsitektur

NO	Hyperparameter
1	Input Image
2	Convolutional Layer
3	Activation
4	Pooling
5	Kernel
6	Padding
7	Stride
8	Optimizer
9	Early Stopping
10	Learning Rate
11	Loss Function
12	Flatten
13	Fully Connected
14	Output

Training Preparation Sebelum ke tahap training, hal yang harus dilakukan adalah *training preparation*. Hal tersebut dilakukan untuk mempersiapkan hal yang harus di setting untuk keperluan training. Adapun hal yang harus di setting antara lain, yaitu memasukkan arsitekturnya atau model, mengatur *criterion* atau *loss function* yang akan digunakan, mengatur optimizer yang akan digunakan, dan mengatur callback atau jumlah epoch secara otomatis. Empat hal tersebut biasa disebut dengan MCOC (*Model, Criterion, Optimizer, Callback*). Pada penelitian ini kami berfokus pada Metode *Optimizer* sesuai dengan judul yaitu perbandingan optimizer *adam*, *adadelata* dan *SGD*.

C.1.Optimisasi Optimasi dapat diartikan sebagai suatu bentuk mengoptimalkan sesuatu hal yang sudah ada, ataupun merancang dan membuat sesuatu secara optimal.

Model *Deep CNN* dilatih dengan memperbarui parameter semua lapisan di

jaringan secara iteratif dan *optimizer* memainkan peran penting di sini. Algoritma *Gradient descent* merupakan pilihan umum untuk melakukan optimasi pada *neural network*. Untuk meminimalkan fungsi tujuan $J(\theta)$, parameter diperbarui dalam arah kebalikan dari gradien fungsi tujuan $J(\theta)$ terhadap parameter. Pada setiap iterasi, *output* yang diinginkan dan *output* yang diprediksi dibandingkan dan kesalahan di *back-propagated*. Satu dari metrik pengukuran kinerja yang paling populer adalah *cross entropy*. *Cross entropy* nilainya mendekati nol, ketika *output* yang diinginkan dan *output* yang diprediksi persis sama disinilah tujuan utama dari setiap teknik optimasi. Teknik optimasi yang akan digunakan pada penelitian ini adalah:

SGD: Ini adalah teknik optimasi untuk memperbarui parameter θ pada setiap langkah waktu t . Aturan pembaruan bobot SGD diberikan dalam Persamaan :

Di mana d_t mewakili gradien dari fungsi tujuan berdasarkan θ pada langkah waktu t dan η adalah kecepatan belajar (*learning rate*).

5. Adadelata: Ini adalah modifikasi dari Adagrad yang mencoba mengurangnya secara monoton menurunkan kecepatan belajar (*learning rate*). Biasanya, Adagrad mengakumulasi semua gradien kuadrat yang lalu untuk memperbarui tingkat pembelajaran, dengan rumus persamaan:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{D_t + \epsilon}} d_t \quad (2)$$

sedangkan Adadelata menggunakan jendela ukuran tetap s untuk mengumpulkan gradien sebelumnya. Di sini, jumlah gradien dianggap sebagai rata-rata kerusak dari gradien persegi sebelumnya. Jadi, rata-rata *running* $R[d^2]_t$ pada waktu langkah t tergantung pada rata-rata sebelumnya dan seperti gradien sekarang yang ditunjukkan pada Persamaan (4)

$$R[d^2]_t = \gamma R[d^2]_{t-1} + (1 - \gamma) d_t^2 \quad (3)$$

Di sini γ adalah konstanta peluruhan dengan nilai 0,9. Sekarang, kita dapat menulis ulang istilah pembaruan Adagrad dari Persamaan (2) menggunakan parameter $\Delta\theta_t$ seperti di bawah ini:

$$\Delta\theta_t = -\frac{\eta}{\sqrt{D_t + \epsilon}} \mathbf{d}_t \quad (4)$$

Di Adadelata, matriks diagonal D_t telah digantikan oleh rata-rata yang menurun di masa lalu gradien kuadrat $R[d^2]_t$ seperti yang diberikan dalam Persamaan (5). Suku penyebut ini disebut *root mean square* (RMS) kesalahan gradien seperti yang ditunjukkan pada Persamaan (6). Demikian pula, *learning rate* η dalam Persamaan (6) diganti dengan kesalahan *root mean square* dari pembaruan parameter $RMS[\Delta\theta]_t$ sebagaimana terlihat dalam Persamaan (7).

$$\Delta\theta_t = -\frac{\eta}{\sqrt{R[d^2]_t + \epsilon}} \mathbf{d}_t \quad (5)$$

$$\Delta\theta_t = -\frac{\eta}{(RMS)[d]_t} \mathbf{d}_t \quad (6)$$

$$RMS[\Delta\theta]_t = \sqrt{R[\Delta\theta^2]_t + \epsilon} \quad (7)$$

dimana $R[\Delta\theta]_t$ mewakili rata-rata pemutakhiran parameter kuadrat secara eksponensial dan didefinisikan dalam Persamaan.

$$R[\Delta\theta^2]_t = \gamma R[\Delta\theta^2]_{t-1} + (1 - \gamma) \Delta\theta_t^2 \quad (8)$$

Karena $RMS[\Delta\theta]_t$ tidak diketahui untuk langkah waktu saat ini, aturan pembaruan sebelumnya dengan $RMS[\Delta\theta]_{t-1}$ dianggap menghasilkan istilah pembaruan Adadelata dalam Persamaan (9) dan yang terakhir Aturan Adadelata diberikan dalam Persamaan (9).

$$\Delta\theta_t = -\frac{(RMS)[\Delta\theta]_{t-1}}{(RMS)[d]_t} \mathbf{d}_t \quad (9)$$

$$\theta_{t+1} = \theta_t + \Delta\theta_t \quad (10)$$

Adam: metode menyimpan rata-rata yang menurun secara eksponensial dari gradien masa lalu (a_t) yang mewakili momen pertama (mean) dan gradien

kuadrat masa lalu (u_t) yang mewakili momen kedua (varians). Gradient a_t dan u_t dihitung sebagai berikut:

$$a_t = \beta_1 a_{t-1} + (1 - \beta_1) d_t \quad (11)$$

$$u_t = \beta_2 u_{t-1} + (1 - \beta_2) d_t^2 \quad (12)$$

Pada a_t dan u_t bias menuju nol ketika tingkat peluruhan sangat kecil (yaitu β_1 dan β_2 adalah mendekati nol). Untuk mengatasi situasi ini, momen pertama dan momen kedua dengan suku koreksi bias dihitung sebagai berikut:

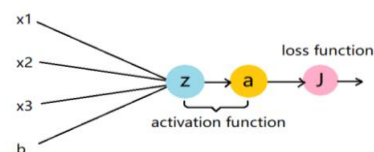
$$\hat{a}_t = \frac{a_t}{1 - \beta_1^t} \quad (13)$$

$$\hat{u}_t = \frac{u_t}{1 - \beta_2^t} \quad (14)$$

Kemudian pembaruan bobot Adam diberikan oleh Persamaan (15).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{u}_t + \epsilon}} \hat{a}_t \quad (15)$$

Rancangan percobaan Pelatihan backpropagation meliputi tiga tahap. Tahap pertama adalah tahap maju (forward). tahap ini menghitung maju tahap layer input sampai tahap layer 29 output dengan menggunakan fungsi aktivasi yang telah ditentukan. Tahap kedua adalah tahap mundur (backward), pada tahap ini selisih antara output jaringan dengan target yang diinginkan merupakan kesalahan yang terjadi. Kesalahan tersebut dipropagasikan mundur, mulai dari garis yang terhubung langsung dengan setiap unit pada layer output. Kemudian tahap yang ketiga adalah tahap yang akan memodifikasi bobot untuk menurunkan tingkat kesalahan yang terjadi (Jumarwanto, 2009). Selain memperbarui fungsi bobot, kesalahan tetap pembaruan harus melewati lapisan di CNN. Sepertinya menggabungkan algoritma Gradient Descent, aturan rantai dan algoritma Back Propagation.



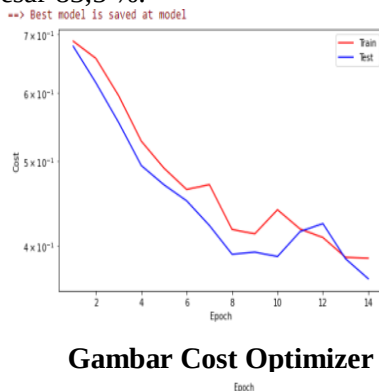
Gambar Backpropagation

adalah diagram sel saraf tunggal. Kami menganggap sampel pelatihan adalah (x, y) , x adalah input, yang melewati fungsi aktivasi $w, b(x)$. Dan unit mendapat output “ a ”, lalu hasilnya “ a ” menuju *loss function*, output dari *loss function* adalah *cost J*.

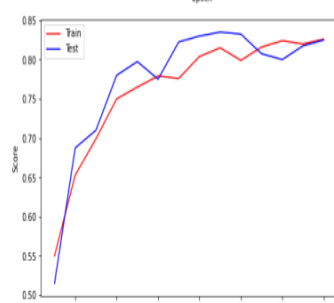
HASIL DAN PEMBAHASAN

Kami telah bereksperimen dengan beberapa model *optimizer* yaitu *Adam*, *Adadelta* dan *SGD* pada model CNN dengan hasil yang dilaporkan pada *tabel III*. Dalam semua percobaan yang telah kami lakukan, kami mendapatkan beberapa hasil akurasi yang berbeda-beda.

Akurasi *Adam* Algoritma optimasi pengganti untuk *stochastic gradient descent* untuk training model deep learning. Adam menggabungkan sifat-sifat terbaik dari algoritma *AdaGrad* dan *RMSProp* untuk memberikan *optimization algorithm* yang dapat menangani *sparse gradients* pada noisy problem [29]. Adam Menggunakan *Lr* 0.01 dan mendapatkan mode akurasi sebesar 83,5 %.



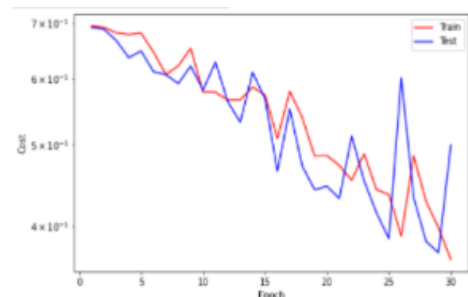
Gambar Cost Optimizer



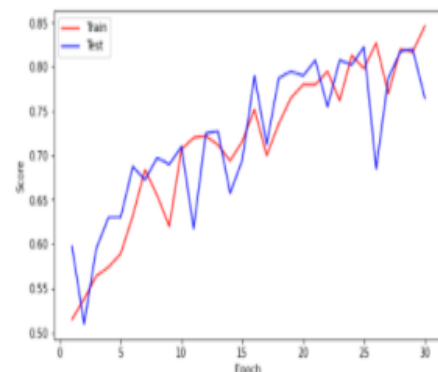
Gambar Plot Score

Dari *gambar 10 dan 11* dapat kita lihat bahwa data test lebih baik dari data training baik dari *cost* dan *score*. Ini berjalan dengan baik.

Akurasi *Adadelta* Metode penurunan *gradien stokastik* yang didasarkan pada tingkat pembelajaran adaptif per dimensi untuk mengatasi [30]. Ini kami menggunakan *Lr* 1,0 dan mendapatkan akurasi sebesar 82,25%.



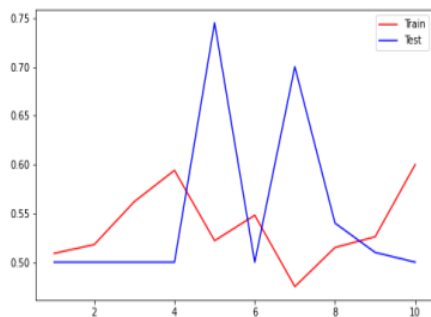
Gambar Hasil Plot Cost Adadelta



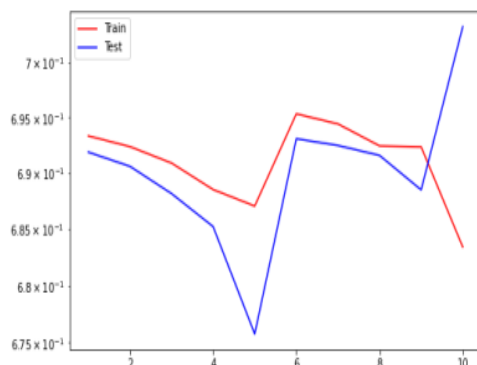
Gambar Hasil Plot Score

Pada *gambar (11) dan (12)* kita bisa melihat jarak permukaan antara data *train* dan *test* sangat bagus yang berpotensi menghindari *over fitting* walaupun akurasinya masih berada dibawah *Adam*.

Stochastic Gradient Descent (SGD) Solusi untuk menyelesaikan issue dari GD. SGD akan melakukan *update weight* tanpa menunggu 1 *epoch* selesai. SGD menggunakan konsep yang mirip dengan *batching* dengan membagi data training menjadi beberapa *batch*. *Weight* akan diupdate dalam setiap *batch* selesai diproses. Dalam metode *optimizer* ini mendapatkan akurasi 69%.



Gambar Hasil Plot Cost SGD



Gambar Hasil Plot Score SGD

kita dapat melihat bahwa pergerakan antara data train dan data test sangat tidak harmonis dan berisiko terjadinya over fitting. Dan metode ini kurang cocok dalam model penelitian yang kita lakukan karena menghasilkan akurasi yang rendah.

Visualisasi Data Untuk melihat bagaimana model bekerja dan apa yang sebenarnya dipelajari, kami memilih untuk memvisualisasikan aktivasi perantara yang terdiri dari hasil fitur map yang dikeluarkan oleh berbagai lapisan konvolusi dan penyatuan dalam jaringan, diberikan input *optimizer* tertentu (keluaran dari suatu lapisan sering disebut aktivasi, output dari fungsi aktivasi). Seperti yang ditunjukkan pada Gambar 16, model belajar bagaimana mengidentifikasi *hydrangea invasive* menggunakan *optimizer AdamW*, terutama pada bunganya.

Tabel Kelas TP,TN,FP,FN

Kelas	Terklasifikasi Positif	Terklasifikasi Negatif
Positif	TP (True Positive)	FN (False Negative)
Negatif	FP (False Positive)	TN (True Negative)

Pada dapat kita lihat bahwa pengukuran kinerja menggunakan confusion matrix. Terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Nilai *True Negative* (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan *False Positive* (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, *True Positive* (TP) merupakan data positif yang terdeteksi benar. *False Negative* (FN) merupakan kebalikan dari *True Positive*, sehingga data positif, namun terdeteksi sebagai data negatif.



Gambar Prediksi

Tabel Eksperimen Hentikan pelatihan saat metrik yang dipantau berhenti meningkat. Dengan asumsi tujuan dari suatu pelatihan adalah untuk

meminimalkan kerugian. Dengan ini, metrik yang akan dipantau adalah 'kerugian', dan modenya adalah 'min'. Loop pelatihan model.fit() akan memeriksa di akhir setiap epoch apakah kerugian tidak lagi berkurang. Setelah ditemukan tidak lagi berkurang, model.stop_training ditandai Benar dan pelatihan dihentikan

Tabel Hasil Eksperimen

Model	Tes
Adam Epoch 14 Train_cost = 0.3869 Test_cost = 0.3663 Train_score = 0.8260 ==> EarlyStop patience = 5 ==> Execute Early Stopping at epoch: 14 Best test_score: 0.8350	Best score: 0.8350
Adadelta Epoch 30 Train_cost = 0.3646 Test_cost = 0.4993 Train_score = 0.8460 Test_score = 0.7650 ==> EarlyStop patience = 5 Best test_score: 0.8225 ==> Execute Early Stopping at epoch: 30	Best score: 0.8225
Epoch 8 Train_cost = 0.6911 Test_cost = 0.6871 Train_score = 0.5260 Test_score = 0.6075 ==> EarlyStop patience = 5 Best test_score: 0.6900 ==> Execute Early Stopping at epoch: 8	Best score: 0.6375

SIMPULAN**DAFTAR PUSTAKA**

Joan G. Ehrenfeld. (2010). "Ecosystem Consequences of Biological

Invasions", Annual Review of Ecology, Evolution, and Systematics, 41: 59–80.

"A neutral terminology to define 'invasive species'", Robert I. Colautti et al., Diversity and Distributions, (Diversity Distrib.) (2004) 10, 135–141.

Kolar, C.S.; D.M. Lodge (2001). "Progress in invasion biology: predicting invaders". Trends in Ecology & Evolution. 16 (4): 199–204.

Reichard, S.H.; C. W. Hamilton (1997). "Predicting invasions of woody plants introduced into North America". Conservation Biology. 11 (1): 193–203.

Williams, J.D.; G. K. Meffe (1998). "Non indigenous Species". Status and Trends of the Nation's Biological Resources. Reston, Virginia: United States Department of the Interior, Geological Survey. 1.

Ewell, J.J.; D.J. O'Dowd; J. Bergelson; C.C. Daehler; C.M. D'Antonio; L.D. Gomez; D.R. Gordon; R.J. Hobbs; A. Holt; K.R. Hopper; C.E. Hughes; M. LaHart; R.R.B. Leakey; W.G. Wong; L.L. Loope; D.H. Lorence; S.M. Louda; A.E. Lugo; P.B. McEvoy; D.M.

Werbos, P.J. (1975). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences.

https://lms.onnocomer.or.id/wiki/index.php/Keras:_Introduction_to_the_Adam_Optimization_Algorithm. Wiki. Retrieved 22 June 201

Belal A . M. Ashqar, Samy S. Abu-Naser, "Identifying Images of Invasive Hydrangea Using Pre-Trained Deep Convolutional Neural Networks", International Journal of Academic Engineering Research (IJAER), Vol. 3 Issue 3, March – 2019

Somenath Bera & Vimal K. Shrivastava, "Analysis of various

- optimizers on deepconvolutional neural network model in the application of hyperspectral remote sensing imageclassification", International Journal of Remote Sensing, VOL. 41, 2019.
- Somenath Bera & Vimal K. Shrivastava." Analysis of various optimizers on deepconvolutional neural network model in the application of hyperspectral remote sensing imageclassification". 2020.
- L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning", 13 December 2017.
- liputan6.com. (2020). serangan hama. Diakses pada 14 januari 2022 dari <https://www.liputan6.com/tag/serangan-hama>
- repository.ipb.ac.id. (2006). Pengaruh penambahan bahan aktif EM4 dan kotoran ayam pada kompos alang-alang, Imperata cylindrica terhadap pertumbuhan semai, Gmelina arborea. Diakses pada 14 januari 2021, dari <https://repository.ipb.ac.id/handle/123456789/46148>
- text-id.123dok.com. Pra-pemrosesan Pre-processing LANDASAN TEORI. Diakses pada 14 januari 2021, dari <https://text-id.123dok.com/document/6zkek81zx-pra-pemrosesan-pre-processing-landasan-teori.html>
- ars.ft.unram.ac.id. (2021). APA ITU ARSITEKTUR. Diakses pada 14 januari 2021, dari <https://ars.ft.unram.ac.id/apa-itu-arsitektur/>
- Max Kuhn and Kjell Johnson." Feature Engineering and Selection: A Practical Approach for Predictive Models", 2019-06-
- wikipedia.org. (1 September 2018). Cara Menulis Daftar Pustaka Dari Website. Diakses pada 14 Januari 2021, dari <https://id.wikipedia.org/wiki/Optimasi>