

Klasifikasi Gambar Gigitan Ular Menggunakan *Regionprops* dan Algoritma *Decision Tree*

Yoga Widi Pamungkas, Adiwijaya*, Dody Qori Utama

Fakultas Informatika, Universitas Telkom, Bandung

Email: ¹yogawidi1@gmail.com, ^{2,*}adiwijaya@telkomuniversity.ac.id, ³dodyqori@telkomuniversity.ac.id

^{*}) Email Penulis Korespondensi: adiwijaya@telkomuniversity.ac.id

Abstrak—Indonesia memiliki biodiversitas hayati ular yang tinggi. Spesies ular yang ada di seluruh wilayah Indonesia, terdiri dari ular berbisa maupun tidak berbisa. Salah satu bahaya yang dapat ditimbulkan oleh ular adalah gigitan beberapa jenis ular yang mematikan. Kasus gigitan ular yang tercatat di Indonesia cukup tinggi dengan korban meninggal dunia yang tidak sedikit. Sebagian kasus kematian yang disebabkan oleh gigitan ular terjadi karena kesalahan dalam prosedur penanganan pada luka gigitannya. Permasalahan ini bisa diatasi salah satunya apabila mengetahui bagaimana cara untuk mengklasifikasikan luka gigitan ular apakah termasuk berbisa atau tidak berbisa. Pada penelitian ini telah dibangun sistem klasifikasi gambar luka gigitan ular menggunakan ekstraksi ciri *Regionprops* dan algoritma *Decision Tree*. Gambar gigitan ular diklasifikasikan termasuk ular berbisa atau tidak berbisa tanpa mengetahui jenis ularnya. Pada *Regionprops* terdapat beberapa fitur yang digunakan untuk membantu proses ekstraksi ciri, antara lain jumlah *centroid*, luas area, *distance*, dan *eccentricity*. Evaluasi dari model yang telah dibangun didapatkan bahwa parameter jumlah *centroid* dan jarak antar *centroid* (*distance*) memiliki pengaruh paling signifikan dalam membantu klasifikasi gambar luka gigitan ular dengan akurasi 97.14%, *precision* 92.85%, *recall* 91.42%, dan *F1 score* 92.06%.

Kata Kunci: *Decision Tree*, *Regionprops*, Gambar Gigitan Ular, *Precision*, *Recall*, *F1 Score*

Abstract—Indonesia has a high biodiversity of snakes. Snake species that exist throughout Indonesia, consisting of venomous and non-venomous snakes. One of the dangers that can be posed by snakes is the bite of several types of deadly snakes. Snake bite cases recorded in Indonesia are quite high with not a few fatalities. Most of the deaths caused by snakebite occur due to errors in the handling procedure for the bite wound. This problem can be overcome one of them if we know how to classify snake bite wounds, whether venomous or non-venomous. In this study, a classification system for snake bite wound image was built using *Regionprops* feature extraction and *Decision Tree* algorithm. Snake bite images are classified as either venomous or non-venomous without knowing the kind of the snake. In *Regionprops* several features are used to help the process of feature extraction, including the number of centroids, area, distance, and eccentricity. Evaluation of the model that was built was found that the parameters of the number of centroids and the distance between centroids had the most significant influence in helping the classification of images of snakebite wounds with an accuracy of 97.14%, precision 92.85%, recall 91.42%, and F1 score 92.06%.

Keywords: *Decision Tree*, *Regionprops*, Snake Bites, *Precision*, *Recall*, *F1 Score*

1. PENDAHULUAN

Indonesia merupakan negara tropis dengan keanekaragaman hayati yang tinggi. Biodiversitas flora yang ada di Indonesia mencapai 20.000 spesies, dengan 40% merupakan tumbuhan endemik Indonesia [1]. Spesies fauna seperti mamalia, burung, reptil, dan ikan mencapai 8.157 jenis [2]. Sekitar 16% spesies amfibi dan reptil yang ada di dunia hidup di Indonesia, dengan jumlah lebih dari 1.100 spesies [3]. Dari sekian banyak spesies reptil di Indonesia, ular adalah salah satu reptil yang dibagi menjadi dua jenis, yaitu berbisa dan tidak berbisa. Spesies ular yang ada di Indonesia sekitar 450 jenis, dengan persebarannya dibagi menjadi dua kelompok berdasarkan asal geografis [4]. Kelompok yang pertama terdiri dari spesies ular yang memiliki kesamaan dengan ular asal Asia yang berada di sebelah barat garis *Wallace* (Sumatra, Jawa, Kalimantan, dan Bali). Kelompok kedua memiliki kesamaan dengan wilayah Australia-Papua. Wilayah ini berada di sebelah timur garis *Wallace* (Papua Barat dan Maluku) [5].

Banyaknya jenis ular yang ada di Indonesia berbanding lurus dengan tingkat kasus gigitan ular yang terjadi. Gigitan ular adalah salah satu permasalahan yang menimbulkan korban jiwa. Perkiraan data kasus gigitan ular yang tercatat di Indonesia sebanyak 113.811 kasus, dengan korban jiwa diperkirakan berjumlah 5.800 [6], [7]. Perkiraan ini berdasarkan laporan penelitian dan mungkin terdapat perbedaan dengan kasus yang sebenarnya. Ada banyak faktor yang mempengaruhi perbedaan ini, seperti kasus gigitan ular yang terjadi di daerah pedesaan, sehingga tidak sampai ke rumah sakit dan tidak ada laporan [8]. Sehingga jumlah kasus gigitan ular mungkin lebih besar dari yang diperkirakan.

Kasus gigitan ular banyak terjadi di kalangan penduduk pedesaan yang miskin, pekerja pertanian, penggembala, nelayan, pemburu, anak-anak yang bekerja (usia 10-14 tahun), penduduk yang memiliki edukasi yang terbatas, dan jauh dari pusat perawatan medis [9]. Dari berbagai kalangan tersebut, sangat minim pengetahuan dalam menangani kasus gigitan ular. Sebagian kasus kematian yang disebabkan oleh gigitan ular terjadi karena kesalahan dalam prosedur penanganan luka pada gigitannya [8]. Penanganan yang dilakukan untuk saat ini masih dilakukan secara tradisional dan manual. Hal ini disebabkan oleh kurangnya teknologi yang dapat digunakan untuk mempermudah klasifikasi gigitan ular. Permasalahan ini bisa diatasi salah satunya dengan mengetahui bagaimana cara untuk mengklasifikasi jenis ular berdasarkan luka gigitan ular.

Penelitian ini bertujuan untuk mengklasifikasikan gambar luka gigitan ular menggunakan *Regionprops* sebagai ekstraksi ciri dan *Decision Tree* sebagai model klasifikasinya. Serta mengetahui parameter mana yang berpengaruh dalam membantu klasifikasi. Sebagai data masukan dari sistem ini adalah gambar bagian tubuh yang hanya terdapat bekas gigitan ular. *Decision Tree* merupakan salah satu metode dalam *Machine Learning* yang mampu divisualisasikan dan mudah untuk dimengerti serta diimplementasikan. *Decision Tree* cocok digunakan pada penelitian ini karena mendukung *biner classification* serta dapat memperoleh hasil yang maksimal walaupun dengan data latih yang tidak terlalu banyak [10]. Sedangkan *regionprops* adalah salah satu ekstraksi fitur dari *image processing* yang dapat mengambil fitur-fitur dari suatu gambar.

Sistem yang dibangun hanya dapat mengklasifikasikan gambar sebagai gigitan ular berbisa atau tidak berbisa tanpa mengetahui jenis dari ular tersebut. Serta dapat mengetahui parameter apa yang paling optimal dari fitur ekstraksi ciri *Regionprops* dalam membantu klasifikasi gambar gigitan ular.

Beberapa penelitian telah dilakukan untuk mengklasifikasikan jenis ular, baik berdasarkan gambar gigitannya maupun gambar dari ular secara langsung. Seperti penelitian oleh Alex J. pada tahun 2017 [11] yang mengklasifikasikan berbagai jenis ular berdasarkan gambar. Metode yang digunakan yaitu klasifikasi k-Nearest Neighbor (kNN) dengan bantuan ekstraksi fitur Scale-Invariant Feature Transform SIFT. Hasil dari penelitian tersebut diketahui bahwa apabila digunakan rasio 0.9 akan memperoleh 634 fitur yang cocok, dengan akurasi yang didapatkan 95,9%. Penelitian lain untuk klasifikasi jenis ular dilakukan oleh Kalana [12]. Penelitian tersebut menguji darah korban apakah terdapat enzim Phospholipase A2 di dalam darah atau tidak. Pada umumnya enzim tersebut terdapat pada ular berbisa.

Penelitian yang dilakukan oleh Rayiemas pada tahun 2018 [13] untuk mengklasifikasikan ular berbisa atau tidak menggunakan ekstraksi fitur *chain code* dan juga kNN. Nilai akurasi yang didapatkan pada penelitian tersebut sebesar 76.92% dengan nilai K adalah 13. Saran yang diberikan pada penelitian tersebut yaitu mengganti metode klasifikasi dan juga menfokuskan pada tahap preprocessing.

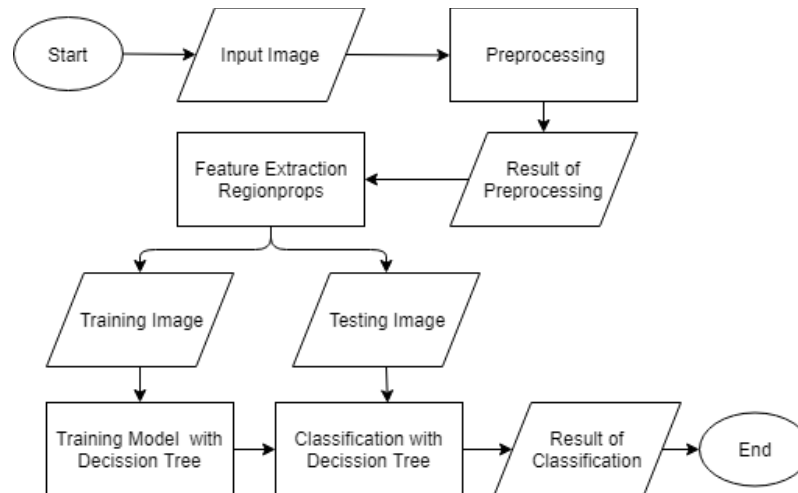
Pada tahun 2019, Hermawati juga membangun sistem yang bisa mengidentifikasi ular berdasarkan gigitannya menggunakan *Local Binary Patern* (LBP) dan juga *Support Vector Machine* (SVM) [14]. LBP digunakan sebagai fitur ekstraksi yang dapat menghasilkan karakteristik dari gambar input dalam bentuk histogram, sedangkan SVM untuk klasifikasinya. Pada penelitian ini dilakukan dua skenario. Skenario yang pertama dengan mengubah parameter point (P) dan radius (r) pada LBP. Sedangkan yang kedua pada parameter C dan sigma pada SVM. Hasil yang didapatkan dengan parameter P = 8, r = 1, C = 1 dan sigma = 0,15 menghasilkan akurasi sebesar 89%.

Aydadenta pada 2018 [15] melakukan klasifikasi menggunakan random forest yang dikombinasikan dengan seleksi fitur berbasis clustering. Akurasi yang diperoleh adalah 98.9% untuk kanker paru, 89% untuk kanker prostat, dan 85.87% untuk kanker usus besar. Amalya melakukan penelitian pada tahun 2018 [16] untuk mendeteksi kanker menggunakan Algoritma *Binary Particle Swarm Optimization* (BPSO) dan *C4.5 Decision Tree*. Dari penelitian yang dilakukan, performansi dari sistem yang dibangun untuk lima data yang digunakan mendapatkan rata-rata akurasi 54% untuk skema klasifikasi dengan Information Gain sebagai seleksi fitur dan rata-rata akurasi 99% untuk skema klasifikasi dengan BPSO sebagai seleksi fitur. Pemilihan nilai k pada diskritisasi didapatkan hasil yang terbaik adalah nilai k = 5.

Shoumy pada tahun 2016 [17] melakukan penelitian tentang fitur ekstraksi untuk mengenali pola pada darah menggunakan *Neural Network*. *Regionprops* adalah fitur yang digunakan untuk mengenali pola objek dengan cara mengekstraksi karakteristik dasar dari noda darah. Di dalam fitur ini terdapat beberapa fitur lagi, diantaranya 'Area', 'Centroid', 'MajorAxisLength', 'MinorAxisLength', dan lainnya. Hasil dari *Regionprops* ini yang akan diolah menggunakan 2 tipe dari Neural Network, yaitu *cascade forward neural network* (CFNN) and *function fitting neural network* (FFNN). Hasil yang diperoleh bahwa FFNN lebih baik daripada CFNN dengan hasil 91.1% dan 86.6%.

2. METODE PENELITIAN

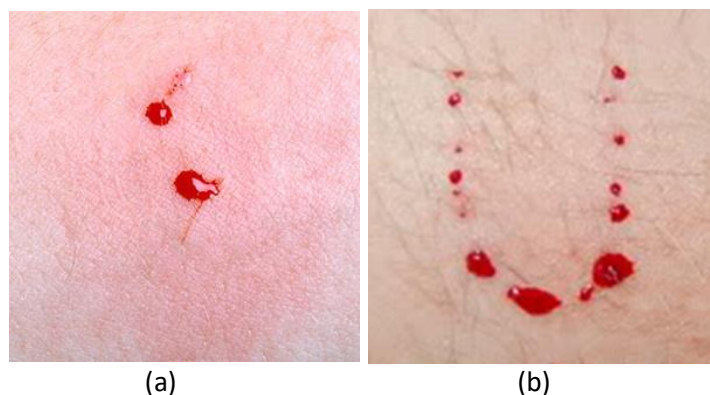
Sistem pada penelitian ini dibagi menjadi beberapa tahap yaitu, tahap preprocessing, tahap ekstraksi ciri, dan tahap pengujian model klasifikasi. Gambar 1 menunjukkan flowchart dari sistem.



Gambar 1. Flowchart dari sistem

2.1 Dataset

Data yang digunakan dalam penelitian ini adalah dataset gambar luka gigitan ular yang terdiri dari 34 gambar gigitan ular dengan 16 gambar ular berbisa dan 18 tidak berbisa. Gambar 2 adalah contoh dari gambar gigitan ular berbisa dan tidak berbisa yang digunakan sebagai input pada penelitian ini.



Gambar 2. Gambar gigitan ular berbisa (a) dan gambar gigitan ular tidak berbisa (b)

2.2 Preprocessing

Tahap pertama yaitu preprocessing. Beberapa diantara gambar input masih memiliki *noise* yang akan mengganggu dalam proses klasifikasi. *Noise* yang dimaksud berupa citra gigitan ular yang tidak jelas, blur pada saat pengambilan gambar, dan terdapat objek yang lain di sekitar luka gigitan. Maka dari itu perlu dilakukan *preprocessing* untuk menangani data tersebut. Pada awalnya gambar diedit terlebih dahulu menggunakan aplikasi pengolah gambar *Photoshop* untuk menghilangkan *noise* yang tidak bisa dihilangkan oleh sistem. Gambar yang digunakan sebagai data memiliki format RGB, yang terdiri dari 3 warna : merah (*red*), hijau (*green*), dan biru (*blue*). Pada proses *preprocessing* dilakukan *thresholding* untuk masing-masing warna dan kemudian digabungkan. Hasil dari *thresholding* ini dalam bentuk *blackwhite* (bw). Karena masih ada bagian-bagian yang belum terisi sempurna pada gambar dilakukan *fill holes* menggunakan fitur *imfill*. Langkah terakhir dari *preprocessing* yaitu menggunakan fitur *strel* ('*disk*', 25) untuk menghilangkan *noise* yang masih tersisa.

2.3 Ekstraksi Ciri *Regionprops*

Hasil dari *preprocessing* selanjutnya masuk ke tahap ekstraksi. Penelitian ini menggunakan *Regionprops* sebagai fitur ekstraksi. Hasil dari ekstraksi ini berupa ciri khusus yang merepresentasikan gambar input. Ada beberapa fitur yang dapat dipakai di dalam *Regionprops*. Fitur yang dipakai pada penelitian ini antara lain, *centroid*, *area*, dan *eccentricity*. *Centroid* merepresentasikan titik bekas luka gigitan ular. *Area* merepresentasikan luas masing-masing *centroid*.

Fitur *Centroid* akan menghasilkan koordinat dari masing-masing objek yang ada pada suatu gambar. Dengan fitur ini dapat mengetahui dua macam parameter, yaitu jumlah objek pada suatu gambar yang merupakan jumlah luka gigitan ular dan juga jarak rata-rata antar objek. Berdasarkan pengamatan, jumlah luka gigitan pada ular berbisa berkisar antara 1 sampai 4 luka, sedangkan untuk ular tidak berbisa meninggalkan bekas luka lebih dari 4. Jarak rata-rata antar objek dihitung menggunakan metrik pengukuran *Euclidean Distance*.

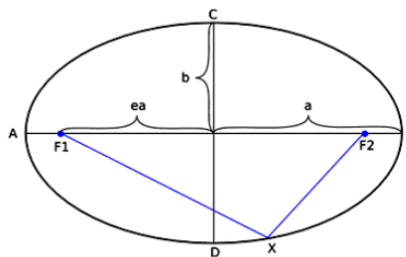
$$d = \sum_{i=1}^{k-1} \sum_{j=1+k}^k \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (1)$$

$$\bar{d} = \frac{d}{k} \quad (2)$$

Pada persamaan (1) digunakan untuk menghitung jarak rata-rata (*distance*) antar setiap objek. Variabel d merupakan jarak antara dua buah objek. Variabel x dan y adalah koordinat suatu objek yang didapatkan dari fitur *centroid*. Variabel k adalah jumlah dari *centroid* yang ada pada gambar. Sedangkan variabel \bar{d} merupakan jarak rata-rata antara semua objek yang ada.

$$A = \sum_{i=1}^k a_i \quad (3)$$

Fitur *area* menghasilkan nilai luas dari masing-masing objek pada gambar. Persamaan (3) untuk menghitung nilai luas total dari semua objek yang ada. Variabel A merepresentasikan luas total dari semua objek. Variabel k adalah jumlah dari objek yang ditemukan pada gambar, a_i merepresentasikan luas objek ke i .



Gambar 3. Ilustrasi *eccentricity*

$$e = \sqrt{1 - \frac{b^2}{a^2}} \quad (4)$$

Fitur *eccentricity* digunakan untuk membedakan bentuk suatu objek dengan lainnya. Nilai perbandingan antara jarak foci ellips minor dengan foci ellips mayor suatu objek. *Eccentricity* memiliki rentang nilai antara 0 hingga 1. Objek yang berbentuk memanjang/mendekati bentuk garis lurus, nilai *eccentricity*-nya mendekati angka 1, sedangkan objek yang berbentuk bulat/lingkaran, nilai *eccentricity*-nya mendekati angka 0. Variabel e merepresentasikan *eccentricity*, variabel a sebagai *major axis*, sedangkan b sebagai *minor axis*.

2.4 Klasifikasi Menggunakan Algoritma *Decision Tree*

Setelah melewati tahap ekstraksi ciri, data yang ada kemudian dipisah menjadi data latih dan data uji. Data latih selanjutnya akan masuk ke dalam tahap klasifikasi menggunakan *Decision Tree*. *Decision Tree* adalah salah satu jenis klasifikasi yang merepresentasikan bentuk dari struktur pohon. Dimana setiap node merepresentasikan atribut, cabang merepresentasikan nilai dari atribut, dan daun merepresentasikan kelas. Pada penelitian ini kedalaman *tree* maksimal yang digunakan adalah empat.

Pada penelitian ini menggunakan algoritma C4.5. Algoritma C4.5 merupakan pengembangan dari algoritma ID3. Secara umum, algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut: (i) Pilih atribut sebagai root, (ii) Buat cabang untuk masing-masing nilai, (iii) Bagi kasus dalam cabang, dan (iv) Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama. Pemilihan atribut sebagai root didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan Persamaan (1) [18].

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} * Entropy(S_i) \quad (5)$$

Dimana S adalah notasi himpunan kasus dan A untuk notasi atribut. Variabel n menyatakan jumlah partisi atribut A dan partisi ke i dinotasikan dengan $|S_i|$. Untuk perhitungan entropi dapat dilihat pada persamaan berikut.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (6)$$

Pada persamaan di atas variabel n mewakili jumlah partisi atribut S . Kemudian p_i menyimpan nilai proporsi dari S_i terhadap S .

Parameter yang digunakan untuk proses pengujian sistem antara lain: jumlah *centroid*, luas area, jarak rata-rata antar objek (*distance*), dan *eccentricity*. Pada tahap pengujian juga digunakan *K-fold validation*. Dimana data akan dibagi menjadi K bagian. Setiap bagiannya akan bergantian berperan sebagai data uji.

2.5 Confusion Matrix

Evaluasi dari kualitas model data diukur dengan membangun *confusion matrix*. Dari *confusion matrix* akan didapatkan nilai akurasi, *precision*, *recall*, dan *F1 score*. Akurasi adalah perbandingan antara kelas prediksi yang sebenarnya dengan kelas prediksi. *Precision* adalah rasio prediksi positif yang dilakukan dengan benar terhadap total prediksi positif. *Recall* menunjukkan rasio prediksi positif yang dilakukan dengan benar terhadap seluruh kelas yang ada. Sedangkan *F1 score* adalah rata-rata dari nilai *precision* dan *recall*.

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

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

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

$$F1\ score = 2 \frac{Recall \times Precision}{R + P} \quad (10)$$

Dimana TP (*True Positive*) adalah jumlah semua diprediksi positif dan benar positif. TN (*True Negatif*) merupakan jumlah yang diprediksi salah. FP (*False Positive*) merupakan kelas yang seharusnya diklasifikasikan positif namun diprediksi negatif. FN (*False Negative*) ketika diprediksi negatif ternyata bernilai positif.

Ada beberapa rumus untuk menghitung akurasi, *precision*, *recall*, dan *F1 score*. Beberapa diantaranya dengan cara *micro-averaging*, *macro-averaging*, dan *weighted averaging*. Karena pada penelitian ini hanya membedakan dua kelas, yaitu ular berbisa dan tidak berbisa, maka menggunakan *macro-averaging*. Perhitungan *macro* dilakukan berdasarkan kelas masing-masing. Berikut adalah rumus untuk menghitung dengan cara *macro-averaging*.

$$Macro - Precision = \frac{Precision1 + Precision2}{2} \quad (11)$$

$$Macro - Recall = \frac{Recall1 + Recall2}{2} \quad (12)$$

$$Macro - F1\ score = 2 \frac{(Macro - Precision) \cdot (Macro - Recall)}{(Macro - Precision) + (Macro - Recall)} \quad (13)$$

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian

Pengujian sistem dilakukan untuk menemukan parameter yang paling optimal dan akurat dalam membantu klasifikasi gambar gigitan ular. Jumlah K yang dipakai pada penelitian ditentukan dengan mencari kombinasi nilai yang paling optimal dari hasil akurasi, *precision*, *recall*, dan *F1 score*. Pada penentuan jumlah K yang akan digunakan, parameter yang dipakai antara lain: jumlah *centroid*, luas area, *distance*, dan *eccentricity*. Tabel 1 menunjukkan hasil penentuan *k-fold*.

Tabel 1. Hasil pengujian *k-fold*

<i>K-fold</i>	Akurasi	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
---------------	---------	------------------	---------------	-----------------

3	94.19	66.66	63.73	65.14
4	91.31	62.55	58.12	60.19
5	93.80	80.0	76.98	78.32
6	91.11	75.0	70.55	72.55
7	94.28	85.71	82.85	84.12
8	91.25	81.25	76.87	78.76

Berdasarkan Tabel 1 diketahui bahwa *k-fold* yang paling optimal adalah 5 dan 7. Pada penelitian kali ini dilakukan dua skenario untuk menentukan parameter yang paling berpengaruh dalam mengklasifikasikan gambar gigitan ular. Skenario pertama menggunakan *5-fold validation*, yang berarti perbandingan proporsi data latih dan data uji adalah 80:20. Skenario kedua menggunakan *7-fold validation*, dengan perbandingan data latih dan data uji 85:15. Tabel 2 dan 3 menunjukkan skenario pengujian yang dilakukan.

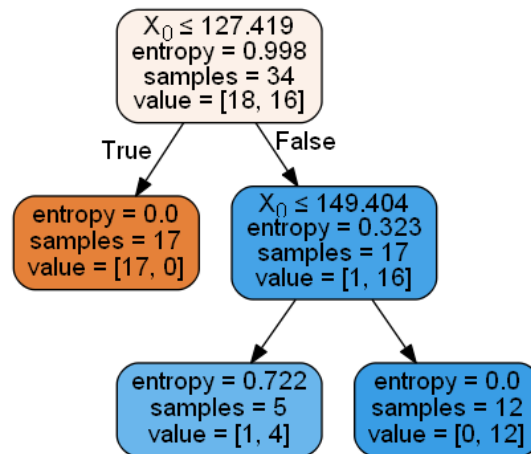
Tabel 2. Hasil pengujian pada skenario pertama dengan *5-fold validation*

Parameter yang digunakan							
Jumlah Centroid	Luas Area	Distance	Eccentricity	Akurasi	Precision	Recall	F1 Score
Ya	-	-	-	93.80	87.55	85.83	86.23
-	Ya	-	-	58.09	51.66	35.00	40.36
-	-	Ya	-	97.14	90.00	88.57	89.23
-	-	-	Ya	90.95	77.55	74.40	75.46
Ya	Ya	-	-	93.80	87.50	85.83	86.23
Ya	-	Ya	-	93.80	80.00	76.90	78.32
Ya	-	-	Ya	93.80	87.50	85.83	86.23
-	Ya	Ya	-	97.14	90.00	88.57	89.23
-	Ya	-	Ya	90.95	77.50	74.40	75.46
-	-	Ya	Ya	97.14	90.00	88.57	89.23
Ya	Ya	Ya	Ya	93.80	80.00	76.90	78.32

Tabel 3. Hasil pengujian pada skenario kedua dengan *7-fold validation*

Parameter yang digunakan							
Jumlah Centroid	Luas Area	Distance	Eccentricity	Akurasi	Precision	Recall	F1 Score
Ya	-	-	-	94.28	90.44	89.04	89.20
-	Ya	-	-	55.00	55.95	37.73	43.01
-	-	Ya	-	97.14	92.85	91.42	92.06
-	-	-	Ya	91.42	83.33	80.47	81.26
Ya	Ya	-	-	94.28	90.47	89.04	89.20
Ya	-	Ya	-	94.28	85.71	82.85	84.12
Ya	-	-	Ya	94.28	90.47	89.04	89.20
-	Ya	Ya	-	97.14	92.85	91.42	92.06
-	Ya	-	Ya	91.42	83.33	80.47	81.26
-	-	Ya	Ya	97.14	92.85	91.42	92.06
Ya	Ya	Ya	Ya	94.28	85.71	82.85	84.12

Pada grafik di bawah ini adalah hasil *tree* yang didapatkan ketika hanya menggunakan parameter *distance*.



Gambar 4. Model *Decision Tree* ketika menggunakan parameter *distance*

3.2 Analisis Hasil Pengujian

Pada pengujian sistem skenario pertama dengan *5-fold validation* didapatkan hasil bahwa parameter *distance* apabila digunakan sendiri mendapatkan akurasi, *precision*, *recall*, dan *F1 score* yang paling tinggi yaitu masing-masing 97.14%, 90.00%, 88.57%, dan 89.23%. Sedangkan yang paling rendah adalah parameter luas area dengan akurasi 58.09%, *precision* 51.66%, *recall* 35.00%, dan *F1 score* 40.36%. Ketika dua parameter digunakan secara bersamaan terbukti bahwa parameter *distance* dapat tetap memperoleh hasil pengujian yang tinggi. Parameter jumlah *centroid* mendapatkan hasil dibawah parameter *distance*, dengan akurasi 93.80%, *precision* 87.55%, *recall* 85.83%, dan *F1 score* 86.23%. Pada Gambar 4 menunjukkan contoh model *tree* yang dihasilkan ketika menggunakan parameter *distance*. Didapatkan hasil bahwa pada *root*, entropi yang didapatkan 0.99. Ketika parameter *distance* dan jumlah *centroid* digunakan secara bersamaan justru menghasilkan nilai yang lebih rendah dibandingkan jika digunakan bersama parameter *eccentricity* dan luas area. Terdapat perbedaan disaat dalam memprediksi suatu gambar yang sama ketika parameter jumlah *centroid* dan *distance* digunakan secara bersamaan.

Skenario kedua menggunakan *7-fold validation* yang secara umum menghasilkan nilai yang agak berbeda dengan skenario pertama. Parameter jumlah *centroid* mengalami peningkatan hasil dengan akurasi 94.28%, *precision* 90.44%, *recall* 89.04%, dan *F1 score* 89.20%. Peningkatan yang lain juga dialami oleh parameter *eccentricity*, sedangkan parameter luas area mengalami penurunan akurasi menjadi 55.00%.

Pada kedua skenario parameter luas area menjadi parameter yang paling tidak optimal. Hal ini terjadi karena parameter luas area sangat bergantung pada saat pengambilan gambar yang digunakan untuk data latih. Semakin dekat jarak pengambilan gambar, semakin besar pula luas area yang dihasilkan. Sehingga parameter luas area kurang optimal sebagai parameter dalam mengklasifikasikan gambar gigitan ular.

5. KESIMPULAN

Pada penelitian ini telah dilakukan serangkaian penelitian untuk menemukan nilai terbaik dalam klasifikasi gambar gigitan ular menggunakan ekstraksi ciri *regionprops* dan *Decision Tree* sebagai algoritma *classifier*. Kombinasi dari beberapa parameter dari *regionprops* telah diterapkan untuk mencari tahu parameter yang paling berpengaruh dalam klasifikasi. Hasil dari penelitian ini diketahui bahwa kombinasi parameter *distance* dengan luas area atau dengan *eccentricity* ketika menggunakan *7-fold validation* menghasilkan nilai yang paling tinggi dengan akurasi 97.14%, *precision* 92.85%, *recall* 91.42%, dan *F1 score* 92.06%.

Parameter jumlah *centroid* dan *distance* sama-sama berpengaruh besar dalam penelitian ini. Apabila digunakan secara terpisah, kedua parameter ini menghasilkan nilai yang tinggi. Namun apabila digunakan secara bersamaan nilai yang dihasilkan menurun dikarenakan terdapat suatu gambar input yang memiliki kualitas kurang bagus, sehingga terdapat perbedaan antara *value* gambar yang asli dengan yang sudah diproses. Kemudian parameter luas area tidak terlalu berpengaruh pada penelitian ini, sehingga apabila digunakan secara bersamaan dengan parameter lain, nilai yang dihasilkan akan mengikuti parameter lain tersebut. Hal ini terjadi karena parameter luas area sangat bergantung pada saat pengambilan gambar yang digunakan untuk data latih. Semakin dekat jarak pengambilan gambar, semakin besar pula luas area yang dihasilkan. Sehingga parameter luas area kurang optimal sebagai parameter dalam mengklasifikasikan gambar gigitan ular.

Beberapa data pada penelitian yang kualitasnya kurang bagus menyebabkan hasil pada tahap *preprocessing* kurang optimal. Pada penelitian mendatang diharapkan menggunakan data yang lebih baik dan juga dilakukan penambahan data. Apabila pada penelitian mendatang dilakukan penambahan data, maka dapat menggunakan model yang lain sebagai klasifikasinya. Kemudian tidak hanya mengklasifikasikan antara ular berbisa dan tidak

berbisa, namun dapat ditambahkan klasifikasi pada jenis-jenis bisa ular, seperti neurotoksin, hemotoksin, dan lainnya.

REFERENCES

- [1] Kusmana, C. (2015). The Biodiversity of Flora in Indonesia. *Bogor Agricultural University*.
- [2] Yuwono, A. (2014). The Fifth National Report Of Indonesia To The Convention On Biological Diversity. *Ministry Of Environment And Forestry Of Indonesia*.
- [3] BAPPENAS (1993). Biodiversity Action Plan for Indonesia. *Ministry of National Development Planning/National Development Planning Agency, Jakarta, Indonesia*. 141 p.
- [4] Latus, P., Panaro, F. (2018). Venomous Snakes – Snakebites. *Introducing Indonesia, 5th edition, AWA*. Available at : <https://www.expat.or.id/medical/venomoussnakes.html>
- [5] Auerbach PS, Norris RL. (2015). Disorders caused by venomous snakebites and marine animal exposures. In: *Harrison's principles of internal medicine*. 18th ed. Vol. 2. New York: McGraw-Hill Professional. p. 3566-76
- [6] Gutiérrez JM, Warrell DA, Williams DJ, et al. (2013). The need for full integration of snakebite envenoming within a global strategy to combat the neglected tropical diseases: the way forward. *PLoS Negl Trop Dis* 7(6):e2162
- [7] International Society on Toxinology. (2014). <http://knoema.com/ueecok/global-snakebite-statistics-may-2014>
- [8] Kasturiratne A, Wickremasinghe AR, de Silva N, et al. (2008). The global burden of snakebite: a literature analysis and modelling based on regional estimates of envenoming and deaths. *PLoS Med*. 5(11):e218.
- [9] World Health Organization (2016). <https://www.who.int/snakebites/epidemiology/en/>
- [10] Gupta, B., Rawat, A., Jain, A., et al. (2017). Analysis of Various Decision Tree Algorithms for Classification in Data Mining. *International Journal of Computer Applications (0975 – 8887)*
- [11] James, A. (2017). Snake classification from images. *PeerJ Preprints* 5.
- [12] Maduwage, K., O'leary, M.A., & Isbister, K.G. (2014) Diagnosis of snake envenomation using a simple phospholipase A2 assay. *Scientific reports* 4.
- [13] Putra, R. M., Adiwijaya & Utama, D.Q. (2018). Snake Bite Classification Using Chain Code and K Nearest Neighbour. *Journal of Physics: Conference Series*.
- [14] Hernawati, N.P.A.U.D, Adiwijaya, & Utama, D. Q. (2018). Image processing for snake identification based on bite using Local Binary Pattern and Support Vector Machine method. *Journal of Physics: Conference Series*, 1192:012007
- [15] Husna Aydadenta, Adiwijaya, (2018), A Clustering Approach for Feature Selection in Microarray Data Classification using Random Forest, *Journal of Information Processing System* 14(5)
- [16] Amalya Citra Pradana1, Adiwijaya2, Annisa Aditsania3. (2018). Implementing binary particle swarm optimization and C4.5 decision tree for cancer detection based on microarray data classification. *Journal of Physics Conference Series* 1192:012014
- [17] Shoumy, N. J., Yaakop, S.N., Ehkan, P., Ali, M. S., & Khatun, S. (2016). Feature extraction for neural network pattern recognition for bloodstain image analysis. *International Journal of Applied Engineering Research*, 11(15):8583-8589.
- [18] Craw, S. (2005). Case Based Reasoning. *Lecture 3 : CBR Case-Base Indexing*. Available at : www.comp.rgu.ac.uk/staff/smc/teaching/cm3016/Lecture-3-cbr-indexing.ppt