

**KLASIFIKASI AUTIS MENGGUNAKAN EKSTRAKSI FITUR
SPEED UP ROBUST FEATURE (SURF)
DENGAN ALGORITMA *BOOSTING***

SKRIPSI

YORRIS SIAGIAN

NPM. 198160014



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MEDAN AREA

2024

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

Document Accepted 6/6/24

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah

3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)6/6/24

**KLASIFIKASI AUTIS MENGGUNAKAN EKSTRAKSI
FITUR *SPEED UP ROBUST FEATURE* (SURF)
DENGAN ALGORITMA *BOOSTING***

SKRIPSI



Diajukan sebagai salah satu syarat untuk
memperoleh Gelar Sarjana (S1) di Fakultas
Teknik Universitas Medan Area

OLEH:

YORRIS SIAGIAN

198160014

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
2024**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

Document Accepted 6/6/24

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah

3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Access From (repository.uma.ac.id)6/6/24

LEMBAR PENGESAHAN

Judul Skripsi : Klasifikasi Autis Menggunakan Ekstraksi Fitur *speed Up Robust Feature (SURF) dengan Algoritma Boosting*

Nama : Yorris Siagian


NPM : 198160014

Fakultas : Teknik

Prodi : Teknik Informatika

Disetujui Oleh

Komisi Pembimbing


Muhathir, ST, M.Kom

Mengetahui:



Eng. Supriatno, ST, MT
Dosen Fakultas Teknik



Rizki Muliono, S.Kom, M.Kom
Kepala Prodi Teknik Informatika

HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun merupakan sebagai syarat untuk memperoleh gelar sarjana dari hasil karya tulis saya sendiri. Adapun bagian-bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain yang telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah. Saya bersedia menerima sanksi akademik yang saya peroleh dan sanksi-sanksi lainnya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.

Medan, 1 April 2024

Yang membuat Pernyataan



Yorris Siagian

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan di bawah ini :

Nama : Yorris Siagian
NPM : 198160014
Program Studi : Teknik Informatika
Fakultas : Teknik
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area **Hak Bebas Royalti (Non-Exclusive Royalti-Free Right)** atas karya ilmiah saya yang berjudul :

Klasifikasi Autis Menggunakan Ekstraksi Fitur *speed Up Robust Feature (SURF)* dengan *Algoritma Boosting*

Dengan Hak Bebas Royalti Non Eksklusif ini Universitas Medan Area berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Medan

Pada tanggal : Maret 2024

Yang menyatakan



(Yorris Siagian)

KATA PENGANTAR

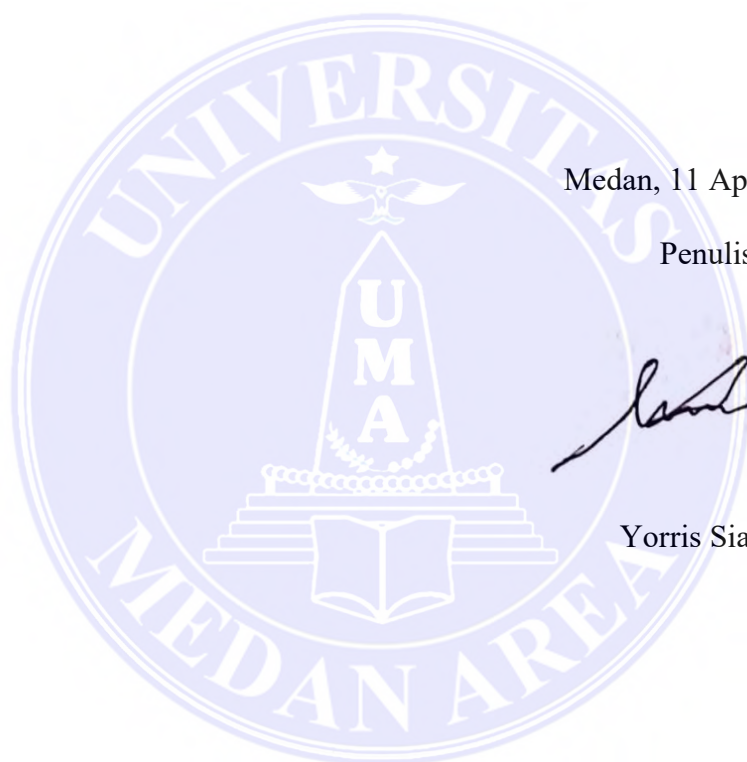
Dengan mengucapkan syukur kepada Tuhan Yang Maha Esa, penulis berhasil menyelesaikan skripsi yang berjudul “Klasifikasi Autistik Menggunakan Ekstraksi Ciri Speed Up Robust Feature (SURF) dengan Algoritma Boosting”, sebagai salah satu syarat untuk memperoleh gelar Sarjana (S1) Program Studi Teknik Informatika Universitas Medan Area. Penyelesaian skripsi ini mendapat dukungan dari berbagai pihak, baik moril maupun materil.

Penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yang Maha Esa, atas kesehatan dan perlindungannya selama masa studi.
2. Keluarga yang selalu memberikan doa, motivasi dan dukungannya yang tiada habisnya.
3. Prof. Dadan Ramdan, M.Eng, M.Sc, Rektor Universitas Medan Area, atas bimbingannya.
4. Dr.Eng. Supriatno, ST, MT, Dekan Fakultas Teknik
5. Ibu Susilawati, S.Kom, M.Kom, Wakil Bidang Penjaminan Mutu Akademik
6. Bapak Rizki Muliono, S.Kom, M.Kom, selaku pimpinan Program Studi Teknik Informatika.
7. Bapak Muhathir, ST, M.Kom, selaku Dosen Pembimbing .
8. Dosen dan Staf Fakultas Teknik Informatika Universitas Medan Area.
9. Rekan-rekan Teknik Informatika angkatan 2019, atas segala pengalaman dan dukungannya.

10. Bundaku Autism Klinik Center yang telah memberikan izin penelitian.

Skripsi ini tidak lepas dari kekurangan, oleh karena itu penulis terbuka menerima kritik dan saran yang bersifat membangun. Harapan penulis semoga skripsi ini dapat memberikan manfaat, baik bagi dunia pendidikan maupun masyarakat luas. Terima kasih.



Medan, 11 April 2024

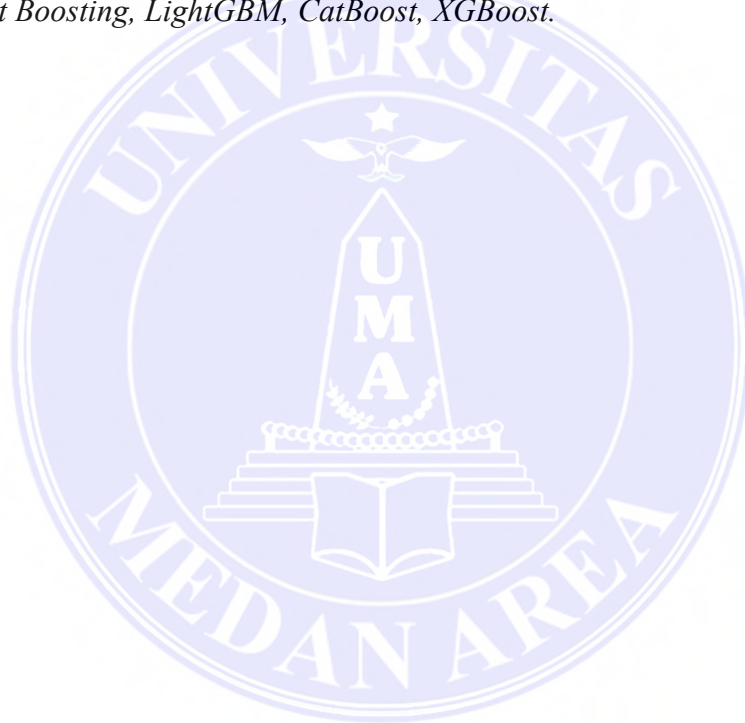
Penulis,

Yorris Siagian

ABSTRAK

Autis adalah gangguan perkembangan neurobehavioral yang kompleks, ditandai dengan pola komunikasi terbatas dan berulang, kesulitan interaksi sosial, serta perbedaan dalam pemrosesan sensori. Penelitian ini bertujuan untuk mengklasifikasikan citra wajah anak autis dan wajah normal menggunakan metode ekstraksi fitur SURF (*Speeded Up Robust Features*) dengan algoritma boosting. Algoritma boosting yang dievaluasi meliputi Adaboost, Gradient Boosting, LightGBM, CatBoost, dan XGBoost. Dataset terdiri dari 102 sampel wajah autis dari bundaku autism clinic center dan 101 sampel wajah normal. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Hasil menunjukkan bahwa algoritma boosting cukup baik dalam mengklasifikasikan wajah autis dan wajah normal. CatBoost mencapai akurasi tertinggi sebesar 80,49%. Hasil ini menunjukkan bahwa algoritma boosting cukup baik dalam klasifikasi wajah autis.

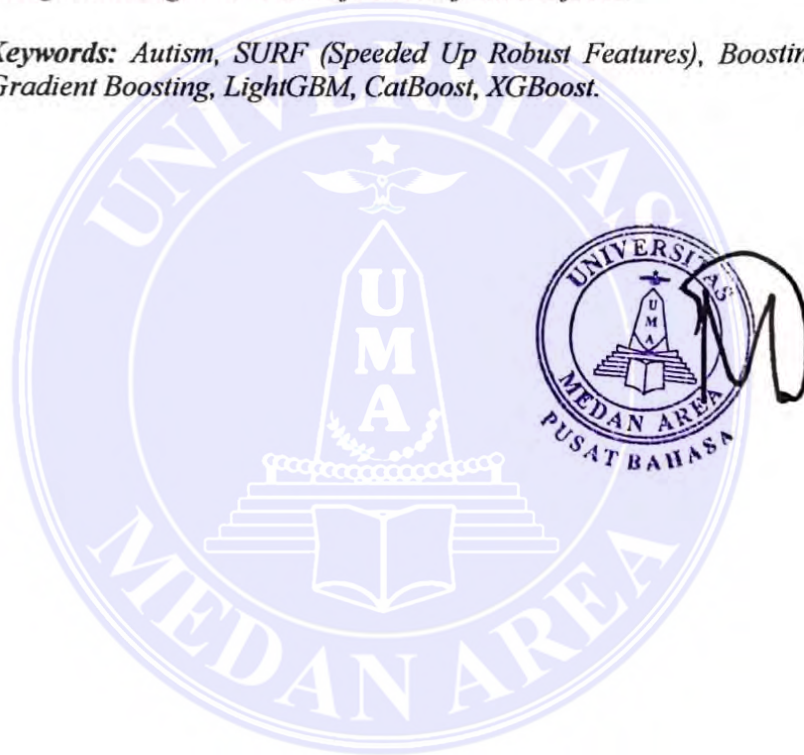
Kata kunci: Autis, SURF (*Speeded Up Robust Features*), *Boosting*, *Adaboost*, *Gradient Boosting*, *LightGBM*, *CatBoost*, *XGBoost*.



ABSTRACT

Autism is a complex neurobehavioral developmental disorder characterized by limited and repetitive communication patterns, social interaction difficulties, and differences in sensory processing. This study aimed to classify facial images of autistic children and normal faces using SURF (Speeded Up Robust Features) feature extraction method with boosting algorithm. The boosting algorithms evaluated include Adaboost, Gradient Boosting, LightGBM, CatBoost, and XGBoost. The dataset consists of 102 autistic face samples from Bundaku autism clinic center and 101 normal face samples. The data was divided into 80% for training and 20% for testing. The results showed that the boosting algorithms were quite good at classifying autistic faces and normal faces. CatBoost achieved the highest accuracy of 80.49%. This result showed that the boosting algorithm was good enough in the classification of autistic faces.

Keywords: *Autism, SURF (Speeded Up Robust Features), Boosting, Adaboost, Gradient Boosting, LightGBM, CatBoost, XGBoost.*



RIWAYAT HIDUP

Penulis dilahirkan di Janjimaria Kabupaten Toba Provinsi Sumatera Utara pada tanggal 07 Maret 2000, dari Ayah Monang Siagian dan Lince Marpaung penulis merupakan anak ke 4 dari 5 bersaudara.

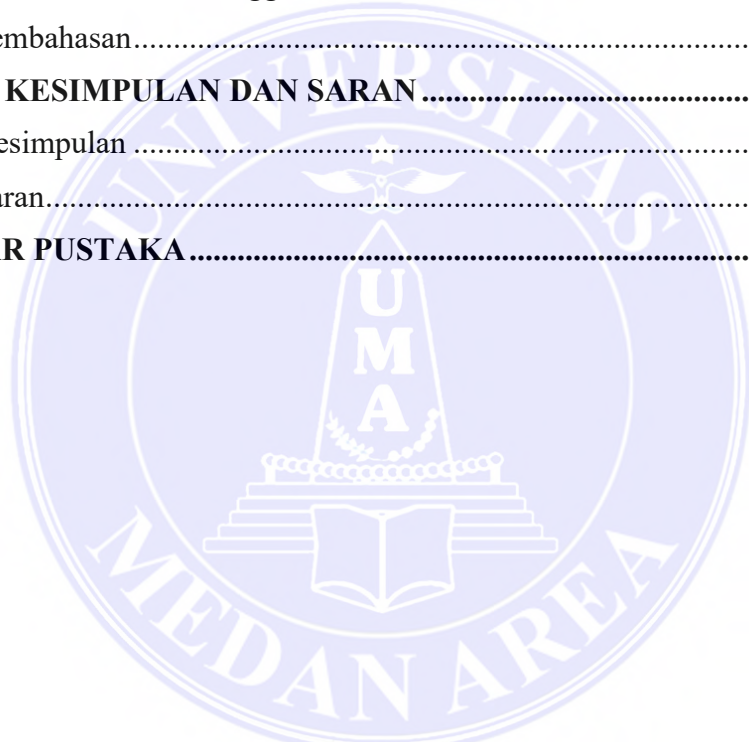
Penulis pertama sekali sekolah di SD Simare-mare pada tahun 2007-2013 dan kemudian melanjutkan Sekolah Menengah Pertama di SMPN 1 Sigumpar lulus pada tahun 2016.

Pada tahun 2017 Penulis lulus dari SMK N 1 Balige dan pada tahun 2019 terdaftar sebagai mahasiswa Fakultas Teknik Prodi Informatika Universitas Medan Area. Penulis mengikuti kerja praktek di PT Del Citra Mandiri selama 1 bulan. Selama proses perkuliahan, penulis aktif berdiskusi belajar terhadap teman-teman dan mahasiswa selain jurusan penulis.

DAFTAR ISI

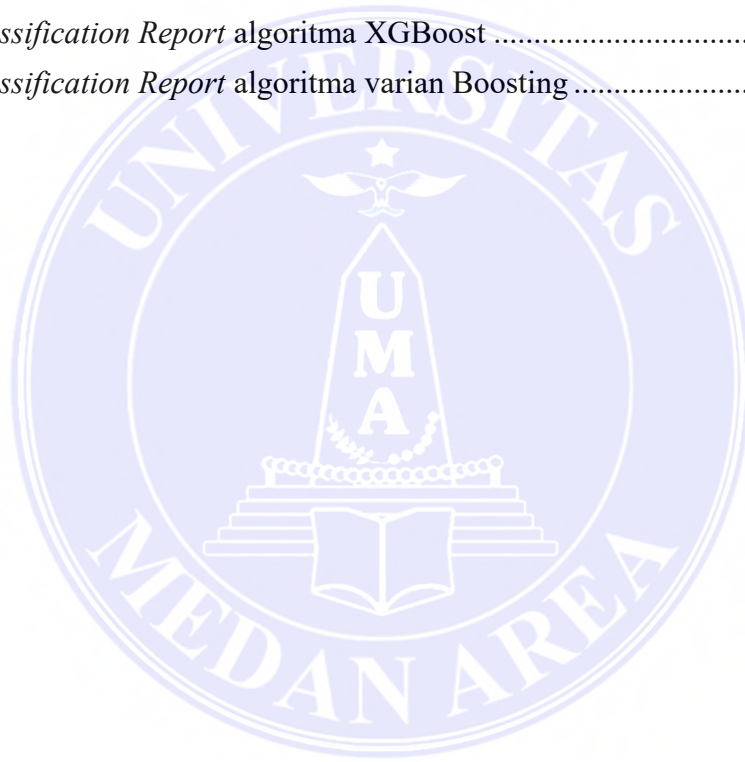
KATA PENGANTAR	iv
ABSTRAK	vi
ABSTRACT	Error! Bookmark not defined.
RIWAYAT HIDUP	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Tujuan	5
1.4. Manfaat Penelitian	5
1.5. Batasan Masalah.....	5
1.6. Sistematika Penulisan	6
BAB II LANDASAN TEORI	7
2.1. Autis	7
2.2. Wajah	7
2.3. Boosting	8
2.3.1. Adaboost (<i>Adaptive Boosting</i>)	8
2.3.2. Gradient Boost (<i>Gradient Boosting</i>)	9
2.3.3. LightGBM (<i>Light Gradient Boosting Machine</i>)	9
2.3.4. Catboost (<i>Categorical Boosting</i>)	10
2.3.5. XGBoost (<i>EXtreme Gradient Boosting</i>)	10
2.4. SURF (<i>Speed Up Robust Feature</i>).....	11
2.5 Metode Evaluasi.....	12
2.6 Penelitian Terdahulu	15
BAB III METODOLOGI PENELITIAN	21
3.1. Alat dan Bahan Penelitian.....	21
3.1.1. Perangkat Keras	21
3.1.2. Perangkat Lunak.....	21
3.2. Metode Pengumpulan data.....	22
3.2.1 Dataset.....	22
3.3 Arsitektur Penelitian	23

BAB IV HASIL DAN PEMBAHASAN	25
4.1 Hasil	25
4.1.1 Informasi Data.....	25
4.1.2 Ekstraksi Fitur <i>SURF</i>	25
4.1.3 Akurasi Hasil Klasifikasi dengan Variasi Algoritma <i>Boosting</i>	26
4.1.3.1 Klasifikasi Menggunakan AdaBoost	27
4.1.3.2 Klasifikasi Menggunakan Gradient Boost	29
4.1.3.3 Klasifikasi Menggunakan LightGBM.....	32
4.1.3.4 Klasifikasi Menggunakan CatBoost.....	34
4.1.3.5 Klasifikasi Menggunakan <i>XGBoost</i>	37
4.2 Pembahasan.....	40
BAB V KESIMPULAN DAN SARAN	43
5.1 Kesimpulan	43
5.2 Saran.....	43
DAFTAR PUSTAKA	44



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	15
Tabel 3.1 Perangkat Keras.....	21
Tabel 3.2 Perangkat Lunak.....	22
Tabel 4. 1 <i>Classification Report</i> algoritma Adaboost	29
Tabel 4. 2 <i>Classification Report</i> algoritma Gradientboost.....	32
Tabel 4. 3 <i>Classification Report</i> algoritma LightGBM	34
Tabel 4. 4 <i>Classification Report</i> algoritma CatBoost.....	37
Tabel 4. 5 <i>Classification Report</i> algoritma XGBoost	40
Tabel 4. 6 <i>Classification Report</i> algoritma varian Boosting	41



DAFTAR GAMBAR

Gambar 2. 1 Confusion Matrix.....	13
Gambar 3.1 Arsitektur Penelitian	24
Gambar 4.1 a. Sampel wajah normal b. Sampel wajah autis.....	25
Gambar 4. 2 a.Hasil ekstraksi fitur wajah normal b. Hasil ekstraksi fitur wajah autis ..	26
Gambar 4. 3 Proses training dan testing adaboost.....	28
Gambar 4. 1 <i>Confusion Matrix</i> dengan Adaboost.....	28
Gambar 4. 5 Proses <i>training</i> dan <i>testing</i> gradient boost.....	30
Gambar 4. 6 <i>Confusion Matrix</i> Gradient Boost.....	31
Gambar 4. 7 Proses <i>training</i> dan <i>testing</i> LightGBM.....	33
Gambar 4. 8 Confusion Matrix dengan LightGBM	34
Gambar 4. 9 Proses <i>training</i> dan <i>testing</i> Catboost	36
Gambar 4. 10 <i>Confusion Matrix</i> dengan CatBoost	37
Gambar 4. 11 Proses training dan testing XGBoost.....	38
Gambar 4. 12 <i>Confusion Matrix</i> dengan XGBoost	39

BAB I

PENDAHULUAN

1.1. Latar Belakang

Autis adalah gangguan kompleks yang mempengaruhi perkembangan *neurobehavioral* dan individu. Gangguan ini ditandai dengan pola komunikasi yang terbatas dan berulang, kesulitan berinteraksi sosial, dan perubahan dalam pemrosesan sensorik (Marotta dkk., 2020). Autis adalah gangguan *neurodevelopmental* yang diwariskan dan kompleks. Gangguan autis mencapai 1% hingga 2% populasi dunia. Biasanya, autis didiagnosis pada usia sekitar 8 tahun dan dapat berbeda tergantung pada faktor seperti jenis kelamin dan lainnya. Kira-kira 25% individu yang mengalami gangguan autis juga memiliki kecacatan intelektual dengan skor IQ di bawah 70 poin (Xie dkk., 2019).

Penelitian autis telah banyak dilakukan sebelumnya seperti dengan mengolah data pencitraan otak besar (Heinsfeld dkk., 2018) menggunakan *Naïve Bayes*, *Support Vector Machine*, *Logistics Regresi*, KNN, *Neural Network* dan *Convolutional Neural Network* pada dataset skrining ASD (*Autism Spectrum Disorder*) anak-anak dan dataset skrining ASD (*Autism Spectrum Disorder*) dewasa (Raj & Masood, 2020) memanfaatkan algoritma *Random Forest*, *Logistic Regression* dan *K-Nearest Neighbors* pada data berbasis kusioner (Abdullah dkk., 2019) menggunakan varian algoritma *naïve bayes* pada data sampel wajah anak normal dan anak autis (Muhathir dkk., 2022).

Berdasarkan capaian penelitian terdahulu penelitian klasifikasi autis telah banyak dilakukan baik dengan data kusioner, data citra otak besar, data wajah dan data skrining ASD. Maka dalam penelitian ini mencoba melanjutkan penelitian dari (Muhathir dkk., 2022) dengan memanfaatkan algoritma *Boosting* dan ekstraksi fitur SURF (*Speed Up Robust Feature*) untuk mengklasifikasi citra wajah autis.

Boosting memiliki keunggulan dalam mengurangi kesalahan algoritma pembelajaran yang lemah, menghasilkan pendekatan dan secara signifikan dan yang secara konsisten menghasilkan klasifikasi yang lebih baik (Tanha dkk., 2020).

Boosting memiliki beberapa metode peningkatan yang banyak digunakan diantaranya Adaboost, Gradient Boost, Catboost XGBoost, dan LightGBM. Adaboost merupakan algoritma klasifikasi yang dimana bobot dari model-model individu disesuaikan berdasarkan kinerjanya, dengan tujuan meningkatkan akurasi keseluruhan dari kumpulan model, dengan memberikan penekanan pada kasus yang salah diklasifikasikan, algoritma ini memungkinkan model-model berikutnya untuk fokus pada kasus yang sulit, sehingga meningkatkan performa klasifikasi. (Sevinç, 2022). Penelitian (Hu dkk., 2020) dengan memanfaatkan algoritma *adaboost* untuk klasifikasi yang bertujuan untuk membedakan pohon pinus yang sakit dan pohon pinus yang sehat dengan mengidentifikasi bayangan dari gambar yang latar belakangnya dihapus. Menunjukkan hasil yang diperoleh AdaBoost lebih baik dibandingkan dengan *K-means clustering*, *support vector machine*, *backpropagation neural networks*, *Alexnet*, *VGG*, and *Inception_v3 networks*.

CatBoost (*Categorical Boosting*) merupakan algoritma pembelajaran mesin yang kuat dan dapat mencapai hasil yang baik dalam berbagai tugas praktis (Dorogush dkk., 2018). Pada penelitian Ibrahim dkk., (2020) dengan menggunakan algoritma *Random forest*, *XGboost*, *Gradient Boost* dan *CatBoost* untuk membandingkan kinerja algoritma dalam pengklasifikasian persetujuan pinjaman dan promosi staf menunjukkan hasil dari algoritma *CatBoost* lebih baik dari pada algoritma lainnya.

XGBoost (eXtreme Gradient Boosting) merupakan salah satu algoritma peningkatan algoritma *boosting* untuk mengklasifikasi pohon regresi model yang berasal dari pohon keputusan peningkatan gradien (Jiang dkk., 2019). Pada penelitian Bentéjac dkk., (2021). Dengan melakukan perbandingan dengan dataset sebanyak 28, dan mendapatkan hasil perbandingan yang menunjukkan *CatBoost* memperoleh hasil terbaik dalam akurasi generalisasi dan AUC, *LightGBM* adalah yang tercepat dibandingkan dengan algoritma *XGBoost* dan *Catboost* tetapi kurang akurat, dan *XGBoost* menempati posisi kedua dalam hal akurasi dan kecepatan latihan.

LightGBM (Light Gradient Boosting Machine) digunakan untuk mengklasifikasi, regresi dan juga pelatihan paralel yang efisien, *LightGBM* juga merupakan jenis *gradient boosting decision tree (GBDT)* dan dianggap jenis *GBDT* yang cepat dan efisien (Shaker dkk., 2021). Pada penelitian Singh dkk., (2020) yang bertujuan untuk prediksi calon pelamar yang akan diterima dengan analisis komparatif menggunakan algoritma *Logistic Regression*, *Neural Networks*, dan *LightGBM* dan hasil yang didapat algoritma *LightGBM* mendapatkan akurasi tertinggi sebesar 98,5%.

Gradient Boost (*Gradient Boosting*) adalah metode ansambel yang umum digunakan untuk meningkatkan performa model, dengan membangun model yang lemah ke arah peningkatan untuk mendapatkan hasil terbaik dalam waktu yang lebih singkat. Metode ini digunakan untuk memecahkan masalah optimalisasi dalam fungsi ruang, dengan prinsip yang mirip dengan metode optimalisasi layaknya peningkatan dalam ruang numerik (J. Wang dkk., 2018)

Pada penelitian ini penulis menggunakan ekstraksi fitur SURF (*Speed Up Robust Feature*) yang merupakan algoritma descriptor untuk deteksi fitur. SURF adalah versi yang telah disempurnakan dari transformasi fitur SIFT (*Scale Invariant Feature Transform*) atau fitur skala invarian, dan bahkan SURF lebih cepat 3-5 kali dari SIFT(Odeh dkk., 2021). SURF banyak digunakan di aplikasi praktis seperti pemahaman adegan dan pengawas serta memiliki kemampuan dalam mengitung operator secara instan dibantu oleh filter kotak, dan SURF memiliki dua tahap yaitu penarikan fitur dan fitur penamaan (Ch & Prasad, 2022).

Berdasarkan kelebihan metode boosting dalam mengklasifikasi data serta kelengkapan informasi yang diekstrak oleh SURF maka, dalam penelitian ini menggunakan metode boosting untuk mengklasifikasikan data citra wajah autisme, dengan memanfaatkan fitur yang diekstrak menggunakan metode SURF. Metode boosting meningkatkan kinerja klasifikasi, sedangkan metode SURF menyediakan fitur citra wajah yang lengkap. Dengan menggabungkan keduanya, penelitian ini bertujuan untuk membangun model komputerasi baru untuk klasifikasi autisme berdasarkan citra wajah.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas maka, rumusan masalah dalam penelitian ini bagaimana mengklasifikasi autis dengan analisis varian algoritma *boosting* dan ekstraksi fitur SURF serta mendapatkan akurasi tertinggi dari varian algoritma *boosting*.

1.3. Tujuan

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah untuk mengklasifikasi autis dengan menggunakan ekstraksi fitur SURF dan varian algoritma *boosting* serta mendapatkan akurasi tertinggi varian algoritma *boosting*.

1.4. Manfaat Penelitian

Berdasarkan tujuan diatas manfaat yang didapat dari penelitian ini adalah untuk mendapatkan hasil klasifikasi autis dari ekstraksi fitur SURF dan akurasi tertinggi varian algoritma *boosting*.

1.5. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Pada penelitian ini, format gambar yang digunakan adalah PNG.
2. Citra yang di *input* merupakan gambar autis yang diambil dengan menggunakan kamera *smartphone*.
3. Proses penelitian ini menggunakan Algoritma varian *boosting*.
4. Implementasi pada penelitian ini menggunakan *Microsoft Visual Studio Code*.

1.6 Sistematika Penulisan

Dalam skripsi ini penulis membahas penggunaan algoritma boosting untuk klasifikasi autis dengan menggunakan ekstraksi fitur SURF yang terbagi menjadi lima bagian utama:

BAB I PENDAHULUAN

Bagian ini membahas tentang latar belakang masalah, tujuan penelitian, manfaat yang diharapkan, dan gambaran struktur penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan teori-teori dasar yang mendukung penelitian ini, termasuk referensi yang digunakan.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang diterapkan dalam penelitian ini, termasuk langkah-langkah yang dilakukan dan hubungannya dengan penelitian yang dilakukan.

BAB IV HASIL DAN ANALISIS

Bagian ini menyajikan hasil pengembangan program dan analisis hasil yang diperoleh.

BAB V KESIMPULAN DAN SARAN

Bab akhir ini menyajikan kesimpulan penelitian dan saran untuk penelitian selanjutnya.

DAFTAR PUSTAKA

Berisi daftar referensi dan sumber yang digunakan selama penelitian, antara lain buku, jurnal, dan artikel, yang disusun menjadi daftar pustaka.

BAB II

LANDASAN TEORI

2.1. Autis

Autis adalah kondisi *neurokognitif* di mana seseorang dengan gangguan autis memiliki kesulitan dalam memahami orang lain, kesulitan berempati, dan cenderung menunjukkan karakteristik yang ekstrem (Botha dkk., 2022). Untuk melakukan diagnosis ASD (*Autism spectrum disorder*) atau sering disebut autis memerlukan proses yang panjang, rumit dan memakan waktu dan juga membutuhkan keahlian klinis yang spesifik. Ada dua instrumen diagnostik yang banyak digunakan yang pertama ADI-R merupakan wawancara semi terstruktur dan diberikan kepada orangtua atau pengasuh yang fokus pada riwayat perkembangan gejala dan yang kedua adalah ADOS merupakan diagnostik observasi skala semi terstruktur standart yang dirancang untuk menilai perilaku sosial serta stereotip dan perilaku yang berulang (Küpper dkk., 2020).

2.2. Wajah

Wajah merupakan salah satu bagian tubuh manusia yang memegang peranan sangat penting, untuk penampilan, ekspresi wajah serta untuk identitas diri sendiri. Tidak ada wajah yang memiliki kembar identik, maupun benar-benar mirip dalam bentuk dan wajah. Pada tubuh manusia wajah berada dibagian depan kepala. Rupa dan bentuk dari wajah manusia diukur berlandaskan pola pada tulang dan otot wajah (Anggraini, 2020).

2.3. Boosting

Boosting merupakan metode yang digunakan untuk meningkatkan akurasi algoritma apapun, yang mengalami *overfitting*, dikarenakan koefisien yang tidak sesuai dengan titik data (Basha dkk., 2018). *Boosting* memiliki beberapa varian diantaranya :

2.3.1. Adaboost (*Adaptive Boosting*)

Adaboost (*Adaptive Boosting*) adalah salah satu algoritma varian *boosting* yang merupakan metode ansambel umum yang menghasilkan kelas yang kuat menggunakan beberapa pengklasifikasi lemah (L. Zhang dkk., 2023).

1. Inisialisasi bobot : $w_i = \frac{1}{N}$.
2. $m = 1 \dots N$.
 - a. Mengambil sampel dari dataset dengan menggunakan bobot $w_i^{(m)}$ untuk mendapatkan sampel pelatihan x_i .
 - b. Melakukan pelatihan pengklasifikasi K_m dengan menggunakan seluruh sampel latihan x_i .
 - c. Menghitung $\epsilon = \frac{\sum_{y_i \neq K_m(x_i)} w_i^{(m)}}{\sum y_i w_i^{(m)}}$, dimana y_i adalah nilai sebenarnya dari variabel target, dan $w_i^{(m)}$ adalah bobot sampel i pada iterasi m .
 - d. Menghitung $\alpha_m = \frac{1}{2} \ln \frac{1-\epsilon}{\epsilon}$.
 - e. Memperbaharui semua bobot $w_i^{(m+1)} = w_i^{(m)} e^{-\alpha_y K_m(x)}$.

$$2. \text{ Prediksi baru dihitung dengan } K(x) = \text{sign} \left[\sum_{m=1}^M \alpha_m K_m(x) \right] \quad (2.1)$$

2.3.2. Gradient Boost (*Gradient Boosting*)

Gradient boosting merupakan suatu algoritma yang melatih sejumlah pembelajar lemah secara berurutan untuk memberikan perkiraan variabel respons dengan tingkat akurasi yang lebih tinggi. Pembelajar lemah adalah model pembelajaran mesin yang memiliki kinerja sedikit lebih baik dari pada kebetulan. Dalam konteks pohon gradient boosting, pembelajar lemahnya biasanya berupa pohon keputusan dangkal. Setiap pohon baru yang ditambahkan ke dalam model ansambel (kombinasi dari semua pohon sebelumnya) dirancang untuk meminimalkan fungsi kerugian yang terkait dengan model ansambel tersebut. Pemilihan fungsi kerugian bergantung pada jenis tugas yang sedang dilakukan dan dapat disesuaikan oleh pengguna. Untuk tugas regresi, fungsi kerugian umumnya adalah kerugian kuadrat. Dengan menambahkan secara berurutan pohon yang meminimalkan fungsi kerugian (mengikuti gradien dari fungsi kerugian keseluruhan), kesalahan prediksi keseluruhan dapat dikurangi. Terdapat banyak hiperparameter yang dapat diatur untuk pohon *gradient boosting*, beberapa mengontrol proses gradient boosting seperti tingkat pembelajaran dan jumlah pohon yang digunakan, sementara yang lain mengatur proses konstruksi pohon, seperti ukuran simpul minimal, sampel dataset yang digunakan, dan kedalaman maksimum (Callens dkk., 2020).

2.3.3. LightGBM (*Light Gradient Boosting Machine*)

LightGBM adalah algoritma pembelajaran mesin berbasis pohon yang cepat dan efisien, dikembangkan oleh *Microsoft* pada tahun 2017 untuk menangani kumpulan data besar dan dimensi tinggi. Pendekatan berbasis histogram digunakan untuk menghitung gradien pada setiap node pohon, dan dapat memanfaatkan teknologi *multi-core* dan GPU.

LightGBM berhasil digunakan dalam berbagai masalah seperti klasifikasi, regresi, dan ranking (D. Zhang & Gong, 2020).

2.3.4. Catboost (*Categorical Boosting*)

Pengklasifikasi CatBoost adalah algoritma pembelajaran mesin varian *boosting* yang memprediksi fitur kategorikal dengan efisien. CatBoost adalah implementasi peningkatan gradien yang menggunakan pohon keputusan biner sebagai prediktor dasar (Ibrahim dkk., 2020).

CatBoost adalah algoritma pembelajaran mesin berbasis Gradient-Boosted Decision Trees (GBDT) dan unggul dalam menangani variabel kategori, meningkatkan akurasi tanpa banyak pra-pemrosesan.

Keunggulan CatBoost mencakup kemampuan mengurangi overfitting dengan mengatasi bias gradien dan kesalahan prediksi. Selama pelatihan, bobot sampel diperbarui secara berurutan berdasarkan hasil putaran sebelumnya, dan hasil akhir diperoleh dengan menggabungkan skor regresi semua pembelajaran yang lemah (Fu dkk., 2024).

2.3.5. XGBoost (*EXtreme Gradient Boosting*)

XGBoost adalah teknik pembelajaran mesin berdasarkan *Gradient Boosting Decision Tree* (GBDT) yang telah banyak digunakan di berbagai bidang data mining. XGBoost menggunakan ansambel pohon regresi untuk membuat prediksi, dan hasil prediksinya adalah penjumlahan dari skor prediksi pohon-pohon dalam model XGBoost

(Bhattacharya dkk., 2020). XGBoost adalah metode peningkatan yang diterapkan pada pohon regresi karena kemampuannya menangani nilai yang hilang (Hamim dkk., 2022).

2.4. SURF (*Speed Up Robust Feature*)

SURF (*Speeded-Up Robust Features*) adalah metode yang digunakan untuk mengenali dan mengekstrak fitur-fitur penting dari gambar secara cepat dan andal. Teknik ini berfokus pada penemuan titik-titik kunci atau poin penting dalam gambar, yang kemudian digunakan untuk menghasilkan deskriptor unik yang mewakili fitur-fitur tersebut.

Proses ekstraksi fitur dengan algoritma SURF melibatkan langkah-langkah berikut:

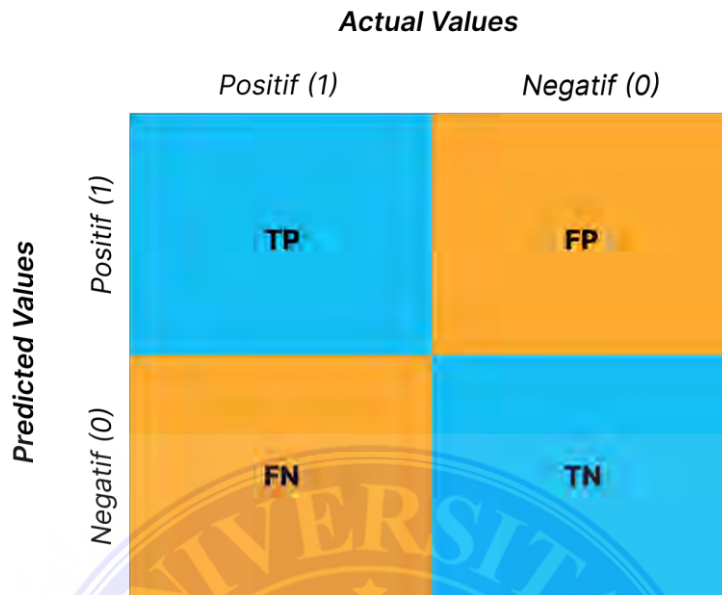
1. Deteksi keypoint cepat: Pada langkah ini, poin-poin penting atau titik-titik kunci diidentifikasi dalam gambar. Algoritma *SURF* menggunakan detektor keypoint cepat invarian lokal untuk menemukan titik-titik penting ini.
2. Penentuan orientasi: Setiap titik kunci diberi orientasi berdasarkan distribusi gradien lokal di sekitarnya. Ini membantu membuat deskriptor menjadi invarian terhadap rotasi.
3. Perhitungan descriptor: Deskriptor unik dihitung untuk setiap titik kunci. Algoritma menggunakan *Haar wavelet* untuk menghitung deskriptor ini. Deskriptor ini mengandung informasi tentang gradien dan distribusi intensitas di sekitar titik kunci.

4. Deskriptor untuk perbandingan: Deskriptor yang dihasilkan dari langkah sebelumnya dapat digunakan untuk membandingkan dan mencocokkan fitur-fitur antara gambar-gambar yang berbeda.

Metode SURF dikenal karena kecepatan komputasinya yang tinggi dan keandalannya dalam mengenali fitur-fitur penting dalam gambar. Dibandingkan dengan teknik ekstraksi fitur SIFT (*Scale-Invariant Feature Transform*) (Gupta dkk., 2021). SURF (*Speed Up Robust Feature*) merupakan descriptor dan pendeteksi gambar yang dikembangkan dari SIFT (*Scale Invariant Feature Transform*) (Nawaz dkk., 2020). Ekstraksi fitur SURF umumnya digunakan untuk pengenalan pola, gambar, klasifikasi maupun rekonstruksi 3D, tetapi ekstraksi fitur *SURF* tidak dapat diterapkan secara langsung untuk gambar multispektral (R. Wang dkk., 2019).

2.5 Metode Evaluasi

Confusion matrix merupakan alat evaluasi yang banyak digunakan dalam pembelajaran mesin untuk klasifikasi atau penentuan perilaku yang diawasi model klasifikasi. Struktur kuadrat dari matriks kebingungan direpresentasikan melalui baris dan kolom, di mana baris adalah kelas aktual dari *instance*, dan kolom adalah kelas yang diprediksi (Hasnain dkk., 2020).



Gambar 2. 1 Confusion Matrix

- TP (*True Positive*) Memprediksi positif dan hasilnya benar.
- TN (*True Negatif*) Memprediksi negative dan hasilnya benar
- FP (*False Positive*) Memprediksi Positif dan hasilnya salah.
- FN (*False Negative*) Memprediksi negatif dan hasilnya salah.

Untuk melakukan klasifikasi ada 5 nilai digunakan, dalam mengukur kemampuan pengklasifikasian yang akan diukur adalah akurasi, presisi, *recall*, *f1 score*, dan *fbeta score*.

a. Akurasi

Akurasi didefinisikan sebagai rasio antara jumlah data yang diklasifikasikan dengan benar (positif yang diprediksi sebagai positif dan negatif yang diprediksi sebagai negatif) dengan jumlah total data dalam dataset (Diantika, 2023).

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN} \dots\dots\dots (2.2)$$

b. Presisi

Presisi adalah evaluasi metrik yang mengukur proporsi sampel yang diprediksi secara akurat positif di antara seluruh sampel yang diprediksi positif (Diantika, 2023).

$$Presicion \frac{TP}{TP+FP} \dots\dots\dots (2.3)$$

c. Recall

Recall adalah proporsi dari jumlah sampel positif yang diprediksi benar dibandingkan dengan seluruh jumlah sampel positif yang sebenarnya ada pada dataset (Diantika, 2023).

$$Recall \frac{TP}{TP+FN} \dots\dots\dots (2.4)$$

d. F1 Score

F1 Score adalah metrik evaluasi yang menggabungkan presisi dan recall untuk memberikan nilai tunggal yang mencerminkan kualitas keseluruhan dari model (Diantika, 2023).

$$F1 \frac{2*Presicion*Recall}{Presicion+Recall} \dots\dots\dots (2.5)$$

e. Fbeta Score

Fbeta Score adalah metrik evaluasi yang sama dengan f1 score, tetapi pada fbeta score memberikan bobot yang lebih besar pada presisi dari pada recall (Le dkk., 2022) dihitung dengan rumus sebagai berikut:

$$f_{\beta}(1 + \beta^2) * \frac{Presicion*Recall}{(\beta^2*Presicion+Recall)} \dots\dots\dots (2.6)$$

2.6 Penelitian Terdahulu

Tabel 2.1 Penelitian Terdahulu

NO	Judul	Metode	Keterangan
1	(Heinsfeld dkk., 2018) “ <i>Identification of autism spectrum disorder using deep learning and the ABIDE dataset</i> ”	SVM, <i>Random Forest</i> , DNN	Untuk mengidentifikasi gangguan spektrum autis pada pasien dari dataset pencitraan otak besar, hanya dengan berdasarkan pola aktivasi otak sipasien. Dicapai akurasi rata-rata klasifikasi 70%, dengan sensitivitas 74% dan spesifisitas 63% dari lipatan validasi silang, dan memiliki akurasi kisaran 66% hingga 71% dalam lipatan individu. Klasifikasi SVM telah mencapai akurasi rata-rata 65% dari 62% menjadi 72%, dengan sensitivitas 68%, dan spesifitas 62%, sementara pengklasifikasi acak telah dicapai akurasi rata-rata 63% dengan sensitivitas 69% dan spesifisitas 58%.

2	(Chen dkk., 2020) <i>“Dissecting Autism Genetic Risk Using Single-cell RNA-seq Data”</i>	A-Risk	Penelitian ini bertujuan untuk mengembangkan metode pembelajaran mesin untuk mengidentifikasi gen risiko autis menggunakan data RNA-seq sel tunggal skala besar dari otak janin manusia. Metode yang dikembangkan adalah A-Risk berbasis gradient boosting, hasil dari penelitian ini adalah pola ekspresi gen risiko yang ditemukan dan kandidat mengungkapkan peran penting neuron rangsang lapisan dalam korteks dewasa dalam etiologi autis. Dengan revolusi transkriptomatik sel tunggal yang belum pernah terjadi sebelumnya dan perluasan kohort autis dengan pengurutan exome atau genom.
3	(Cho dkk., 2019) <i>“Automatic detection of Autism Spectrum</i>	<i>Gradient Boosting</i>	Penelitian ini mendeteksi ASD (<i>Autism Spectrum Disorder</i>) dengan fitur akustik dan teks dari

	<p><i>Disorder in children using acoustic and text features from brief natural conversations”.</i></p>		<p>percakapan singkat dan tidak terstruktur dengan konfederasi. Pada penelitian ini memiliki tujuh puluh anak (35 anak ASD dan 35 anak TD) dengan membahas berbagai topik umum (misalnya, hewan peliharaan, keluarga, hobi, dan olahraga) dengan konfederasi selama sekitar 5 menit. Sebanyak 624 fitur (352 akustik + 272 teks) dimasukkan ke dalam <i>Gradient Boosting</i>. Hasil dari penelitian ini untuk mengklasifikasi identifikasi status diagnostik anak-anak sebesar 75,71%. Untuk anak TD mendapatkan akurasi sebesar 85,71% (30 dari 35 anak) untuk ASD mengalami penurunan sebesar 65,71% (23 dari 35 anak).</p>
4	<p>(Sujatha dkk., 2021) “<i>A Machine Learning Way to Classify Autism Spectrum Disorder</i>”</p>	<p>SVM, KNN, <i>Random Forest</i>, <i>Naive Bayes</i>, <i>Stochastic</i></p>	<p>Tujuan dari penelitian ini adalah untuk mengidentifikasi dengan beberapa metode machine learning untuk menemukan teknik terbaik</p>

		<p><i>gradient descent, adaptive boosting, dan CN2 Rule Induction</i></p>	<p>untuk klasifikasi ASD dari SVM, KNN, <i>Random Forest, Naïve Bayes, Stochastic gradient Descent (SGD), Adaptive Boosting (AdaBoost), dan CN2 Rule</i> dengan menggunakan dataset dari repositori UCI yang diantaranya dataset dewasa, remaja, anak dan balita. Akurasi Hasil klasifikasi yang peneliti peroleh setelah ion eksperimen dalam kasus himpunan data dewasa SGD memberikan 99,7%, dalam himpunan data remaja RF memberikan 97,2%, dalam himpunan data anak SGD memberikan 99,6%, dalam himpunan data balita <i>AdaBoost</i> memberikan 99,8%.</p>
5	<p>(Varshini & Chinnaiyan, 2020) “<i>Optimized Machine Learning Classification Approaches for</i></p>	<p><i>logistic regression, KNN dan Random Forest</i></p>	<p>Penelitian ini bertujuan untuk mengembangkan model klasifikasi menggunakan algoritma pembelajaran mesin untuk mendukung diagnosis dini autis</p>

	<i>Prediction of Autism Spectrum Disorder”</i>		<p>pada anak-anak. Dalam penelitian ini, algoritma pembelajaran mesin <i>random forest</i>, <i>regresi logistik</i>, dan <i>K-Nearest Neighbor</i> (KNN) digunakan dengan dua dataset, yaitu dewasa dan balita. Evaluasi dilakukan untuk mengevaluasi performa teknik klasifikasi pembelajaran mesin dan efisiensi algoritma klasifikasi pada dataset. Hasil yang didapat menunjukkan bahwa algoritma KNN memiliki skor akurasi yang lebih tinggi, yaitu 69,2%, dibandingkan dengan dua algoritma lainnya.</p>
6	(Muhathir dkk., 2022) <i>“Image Classification of Autism Spectrum Disorder Children Using Naïve Bayes Method With Hog Feature Extraction.”</i>	<i>Naive Bayes</i>	<p>Penelitian ini menggunakan variasi metode <i>Naive Bayes</i> dan dibantu dengan ekstraksi fitur <i>Histogram of Oriented Gradient</i> (HOG), dataset yang digunakan diambil langsung dari beberapa SLB (Sekolah Luar Biasa) yang ada dikota Medan.</p>

			Mendapatkan hasil akurasi tertinggi sebesar 89,79%.
--	--	--	---



BAB III

METODOLOGI PENELITIAN

3.1. Alat dan Bahan Penelitian

Pada penelitian ini, penulis menggunakan alat dan bahan untuk mendukung proses penelitian ini, adapun alat dan bahan yang digunakan adalah sebagai berikut:

3.1.1. Perangkat Keras

Berikut adalah perangkat keras yang digunakan dalam penelitian ini, yang dapat membantu proses komputasi dalam memastikan kinerja yang optimal dalam mengklasifikasi autis.

Tabel 3.1 Perangkat Keras

No	Perangkat keras	Keterangan
1	Device	Laptop Lenovo IdeaPad 3 4IML05
2	Processor	Intel(R) Core (TM) i3-10110U CPU @ 2.10GHz 2.59 GHz
3	SSD	237 GB
4	RAM	4,00 GB

3.1.2. Perangkat Lunak

Berikut adalah perangkat lunak yang digunakan dalam penelitian ini, untuk mengolah data dan membantu memaksimalkan pengembangan model dalam mengklasifikasi autis.

Tabel 3.2 Perangkat Lunak

No	Perangkat Lunak	Keterangan
1	Sistem Operasi	Windows 11 Home Single Language
2	Microsoft Visual Studio Code	Version 1.84.2

3.2. Metode Pengumpulan data

Berikut ini adalah metode yang digunakan untuk mengumpulkan dataset yang sangat rinci. Bundaku *autism clinic center*, sebuah lembaga yang berfokus pada penelitian dan perawatan autis, mengumpulkan secara khusus data gambar wajah untuk dataset ini. Beberapa langkah penting dalam pengumpulan data ini termasuk mendapatkan persetujuan dan kerjasama dari pihak klinik, serta masalah etis terkait privasi dan penggunaan data pasien. Selanjutnya, data disimpan dan diorganisasikan dalam bentuk yang memungkinkan analisis lanjutan dan pembuatan model klasifikasi.

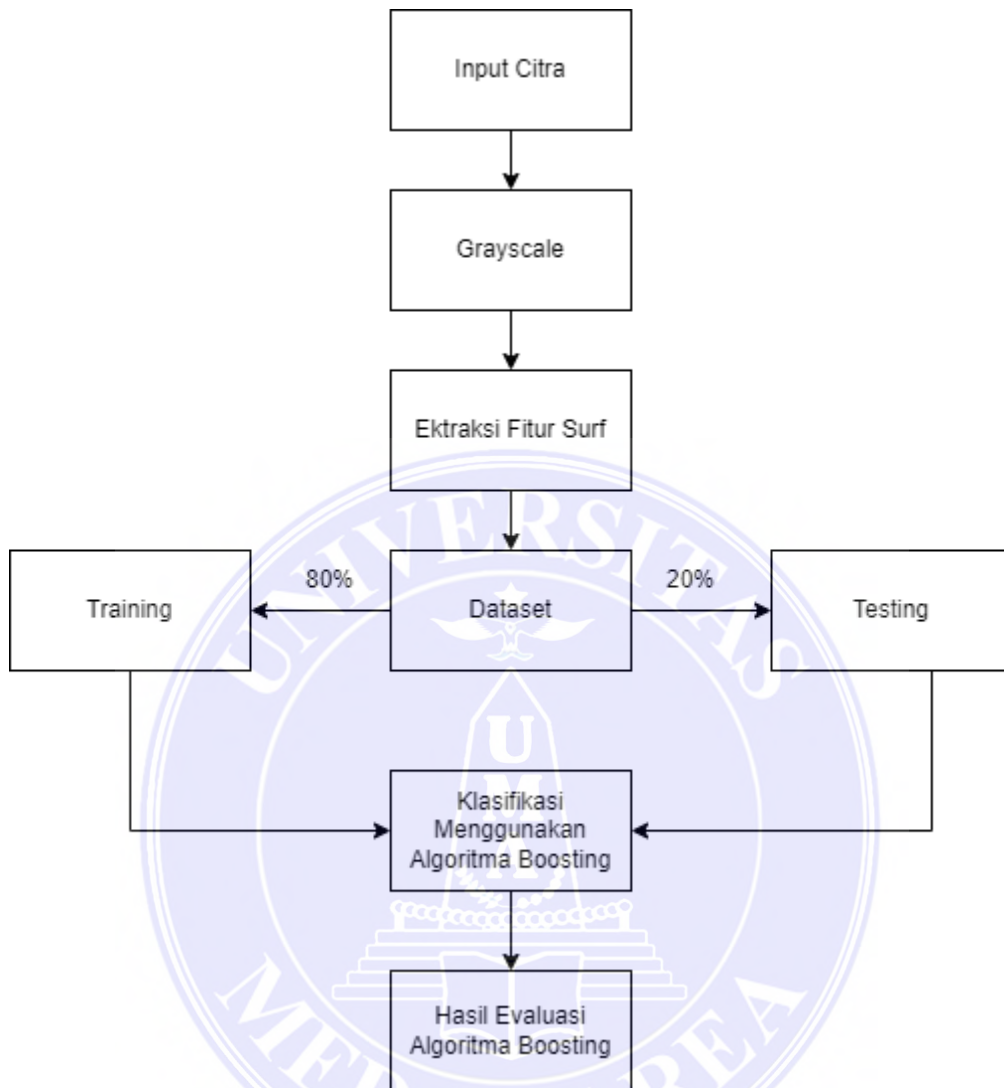
3.2.1 Dataset

Dataset terdiri dari sejumlah citra wajah yang memenuhi kriteria inklusi tertentu. Data yang digunakan berasal dari pasien yang telah didiagnosis dengan gangguan autis oleh profesional medis di bundaku *autism clinic center*. Data sampel yang digunakan sebanyak 102 sampel wajah anak autis dan 101 wajah normal. Model Evaluasi yang digunakan untuk penelitian ini dengan membagi data *training* dan data *testing*, 80% sampel data dijadikan sebagai data *training* sedangkan untuk 20 % sampel data dijadikan sebagai data *testing*. Hal ini sesuai dengan evaluasi model yang digunakan untuk menguji

performa algoritma pada penelitian (Muhathir dkk., 2022). Dengan menggunakan simulasi data *training* dan data *testing* seperti ini, diharapkan hasil evaluasi model dapat lebih akurat dan handal.

3.3 Arsitektur Penelitian

Pada gambar 3.2 adalah proses klasifikasi wajah autis dimulai dengan memasukkan citra wajah autis dan wajah normal sebagai input. Langkah pertama adalah mengubah citra tersebut menjadi grayscale, yang bertujuan mengurangi kompleksitas data dengan menghilangkan informasi warna sehingga fokus pada tekstur dan bentuk yang penting lebih jelas terlihat. Setelah itu, dilakukan ekstraksi fitur menggunakan metode Speeded Up Robust Features (SURF) untuk menyeleksi fitur-fitur penting pada wajah yang diperlukan untuk proses klasifikasi. Data hasil ekstraksi fitur kemudian dibagi menjadi dua bagian, yakni 80% untuk pelatihan (*training*) dan 20% untuk pengujian (*testing*), guna memastikan bahwa model dapat dilatih dengan data yang memadai serta diuji dengan data yang tidak pernah dilihat sebelumnya. Pada tahap pelatihan, data *training* digunakan untuk melatih model dengan algoritma *boosting*, yang bekerja dengan menggabungkan beberapa model lemah menjadi satu model kuat untuk meningkatkan akurasi klasifikasi. Setelah model dilatih, dilakukan pengujian menggunakan data *testing* untuk mengevaluasi kinerja model dalam mengklasifikasikan data baru. Hasil evaluasi algoritma *boosting*, yang merupakan keluaran klasifikasi untuk membedakan wajah autis dan wajah normal, dievaluasi berdasarkan *confusion matrix*, mencakup nilai-nilai seperti akurasi, presisi, *recall*, *f1-score*, dan *fbeta-score* untuk mengetahui seberapa baik algoritma *boosting* dalam melakukan klasifikasi.



Gambar 3.1 Arsitektur Penelitian

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini bertujuan untuk mengklasifikasikan wajah autis dan wajah normal dengan menggunakan algoritma *boosting* dan ekstraksi fitur SURF. Algoritma *boosting* dipilih berdasarkan kemampuannya yang telah terbukti dalam meningkatkan kinerja klasifikasi, sedangkan SURF digunakan karena efektivitasnya dalam menghasilkan fitur citra wajah yang komprehensif. Berbagai metode *boosting* seperti Adaboost, Gradient Boost, LightGBM, Catboost, dan XGBoost telah diuji. Model evaluasi dilakukan dengan membagi data *training* dan *testing*. Sebelum dilakukan optimasi, CatBoost mendapat akurasi tertinggi sebesar 80,49%, kemudian Gradient Boost dan LightGBM yang masing-masing memperoleh akurasi sebesar 73,17%, XGBoost sebesar 70,73%, dan Adaboost sebesar 68,29%.

5.2 Saran

1. Optimalisasi parameter *boosting*, fokus pada penyesuaian *learning rate* dan jumlah *estimator* pada algoritma *boosting* untuk meningkatkan efektivitas klasifikasi autis.
2. Memperbanyak dataset, menambah kumpulan data untuk meningkatkan akurasi model.
3. Peningkatan ekstraksi fitur SURF Mencari metode selain ekstraksi fitur SURF yang lebih efisien atau teknik alternatif untuk meningkatkan kinerja klasifikasi.

DAFTAR PUSTAKA

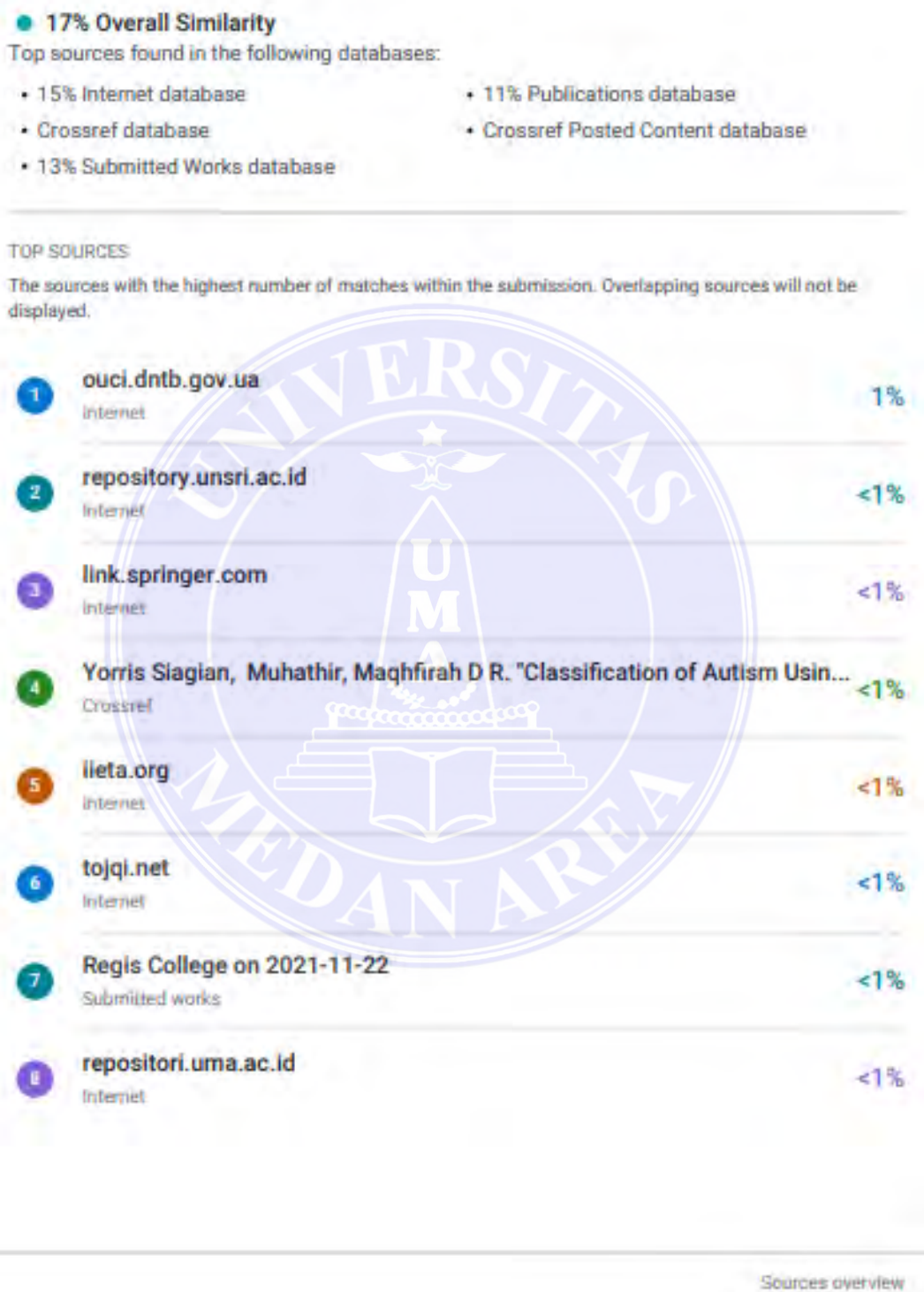
- Abdullah, A. A., Rijal, S., & Dash, S. R. (2019). Evaluation on Machine Learning Algorithms for Classification of Autism Spectrum Disorder (ASD). *Journal of Physics: Conference Series*, 1372(1). <https://doi.org/10.1088/1742-6596/1372/1/012052>
- Anggraini, W. (2020). Deep Learning Untuk Deteksi Wajah yang Berhijab Menggunakan Algoritma Convolutional Neural Network (CNN) Dengan Tensorflow. *Sustainability (Switzerland)*, 4(1).
- Basha, S. M., Rajput, D. S., & Vandhan, V. (2018). Impact of Gradient ascent and Boosting algorithm in classification. *International Journal of Intelligent Engineering and Systems*, 11(1). <https://doi.org/10.22266/ijies2018.0228.05>
- Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3). <https://doi.org/10.1007/s10462-020-09896-5>
- Bhattacharya, S., Siva Ramakrishnan, S., Praveen Kumar Reddy, M., Kaluri, R., Singh, S., Thippa Reddy, G., Alazab, M., & Tariq, U. (2020). A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU. *Electronics (Switzerland)*, 9(2). <https://doi.org/10.3390/electronics9020219>
- Botha, M., Dibb, B., & Frost, D. M. (2022). “Autism is me”: an investigation of how autistic individuals make sense of autism and stigma. *Disability and Society*, 37(3), 427–453. <https://doi.org/10.1080/09687599.2020.1822782>
- Callens, A., Morichon, D., Abadie, S., Delpy, M., & Liquet, B. (2020). Using Random forest and Gradient boosting trees to improve wave forecast at a specific location. *Applied Ocean Research*, 104. <https://doi.org/10.1016/j.apor.2020.102339>
- Ch, L. K., & Prasad, K. V. (2022). *Optimized deep learning model for spatio-temporal detection and localization of object removal video forgery with multiple feature extraction*. <https://doi.org/10.21203/rs.3.rs-1641193/v1>
- Chen, S., Zhou, X., Byington, E., Bruce, S. L., Zhang, H., & Shen, Y. (2020). Dissecting Autism Genetic Risk Using Single-cell RNA-seq Data. *bioRxiv*.
- Chen, Y., Chang, R., & Guo, J. (2021). Emotion Recognition of EEG Signals Based on the Ensemble Learning Method: AdaBoost. *Mathematical Problems in Engineering*, 2021. <https://doi.org/10.1155/2021/8896062>
- Cho, S., Liberman, M., Ryant, N., Cola, M., Schultz, R. T., & Parish-Morris, J. (2019). Automatic Detection of Autism Spectrum Disorder in Children Using Acoustic and Text Features from Brief Natural Conversations. *Interspeech 2019, 2019-September*, 2513–2517. <https://doi.org/10.21437/Interspeech.2019-1452>

- Diantika, S. (2023). Penerapan Teknik Oversampling Untuk Mengatasi Imbalance Class Dalam Klasifikasi Website Phishing Menggunakan Algoritma LightGBM. Dalam *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Nomor 1).
- Dorogush, A. V., Ershov, V., & Gulin, A. (2018). *CatBoost: gradient boosting with categorical features support*. <http://arxiv.org/abs/1810.11363>
- Fu, P., Zhang, J., Yuan, Z., Feng, J., Zhang, Y., Meng, F., & Zhou, S. (2024). Estimating the Heavy Metal Contents in Entisols from a Mining Area Based on Improved Spectral Indices and Catboost. *Sensors*, 24(5). <https://doi.org/10.3390/s24051492>
- Gupta, S., Thakur, K., & Kumar, M. (2021). 2D-human face recognition using SIFT and SURF descriptors of face's feature regions. *Visual Computer*, 37(3). <https://doi.org/10.1007/s00371-020-01814-8>
- Hamim, T., Benabbou, F., & Sael, N. (2022). Student profile modeling using boosting algorithms. *International Journal of Web-Based Learning and Teaching Technologies*, 17(5). <https://doi.org/10.4018/IJWLTT.20220901.0a4>
- Hasnain, M., Pasha, M. F., Ghani, I., Imran, M., Alzahrani, M. Y., & Budiarto, R. (2020). Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. *IEEE Access*, 8, 90847–90861. <https://doi.org/10.1109/ACCESS.2020.2994222>
- Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., & Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, 17, 16–23. <https://doi.org/10.1016/j.nicl.2017.08.017>
- Hu, G., Yin, C., Wan, M., Zhang, Y., & Fang, Y. (2020). Recognition of diseased Pinus trees in UAV images using deep learning and AdaBoost classifier. *Biosystems Engineering*, 194, 138–151. <https://doi.org/10.1016/j.biosystemseng.2020.03.021>
- Ibrahim, A. A., Ridwan, R. L., Muhammed, M. M., Abdulaziz, R. O., & Saheed, G. A. (2020). Comparison of the CatBoost Classifier with other Machine Learning Methods. Dalam *IJACSA International Journal of Advanced Computer Science and Applications* (Vol. 11, Nomor 11). www.ijacsa.thesai.org
- Jiang, Y., Tong, G., Yin, H., & Xiong, N. (2019). A Pedestrian Detection Method Based on Genetic Algorithm for Optimize XGBoost Training Parameters. *IEEE Access*, 7, 118310–118321. <https://doi.org/10.1109/ACCESS.2019.2936454>
- Küpper, C., Stroth, S., Wolff, N., Hauck, F., Kliewer, N., Schad-Hansjosten, T., Kamp-Becker, I., Poustka, L., Roessner, V., Schultebrucks, K., & Roepke, S. (2020). Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learning. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-61607-w>

- Le, H., Peng, B., Uy, J., Carrillo, D., Zhang, Y., Aevertmann, B. D., & Scheuermann, R. H. (2022). Machine learning for cell type classification from single nucleus RNA sequencing data. *PLoS ONE*, *17*(9 September). <https://doi.org/10.1371/journal.pone.0275070>
- Marotta, R., Risoleo, M. C., Messina, G., Parisi, L., Carotenuto, M., Vetri, L., & Roccella, M. (2020). The neurochemistry of autism. *Brain Sciences*, *10*(3). <https://doi.org/10.3390/brainsci10030163>
- Muhathir, M., Muliono, R., & Hafni, M. (2022). Image Classification of Autism Spectrum Disorder Children Using Naïve Bayes Method With Hog Feature Extraction. *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, *5*(2), 494–501. <https://doi.org/10.31289/jite.v5i2.6365>
- Nawaz, S. A., Li, J., Bhatti, U. A., Mehmood, A., Shoukat, M. U., & Bhatti, M. A. (2020). Advance hybrid medical watermarking algorithm using speeded up robust features and discrete cosine transform. *PLoS ONE*, *15*(6). <https://doi.org/10.1371/journal.pone.0232902>
- Odeh, N., Toma, A., Mohammed, F., Dama, Y., Oshaibi, F., & Shaar, M. (2021). An efficient system for automatic blood type determination based on image matching techniques. *Applied Sciences (Switzerland)*, *11*(11). <https://doi.org/10.3390/app11115225>
- Raj, S., & Masood, S. (2020). Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques. *Procedia Computer Science*, *167*, 994–1004. <https://doi.org/10.1016/j.procs.2020.03.399>
- Sevinç, E. (2022). An empowered AdaBoost algorithm implementation: A COVID-19 dataset study. *Computers & Industrial Engineering*, *165*, 107912. <https://doi.org/10.1016/j.cie.2021.107912>
- Shaker, B., Yu, M. S., Song, J. S., Ahn, S., Ryu, J. Y., Oh, K. S., & Na, D. (2021). LightBBB: Computational prediction model of blood-brain-barrier penetration based on LightGBM. *Bioinformatics*, *37*(8). <https://doi.org/10.1093/bioinformatics/btaa918>
- Singh, S. P., Singh, P., & Mishra, A. (2020). Predicting Potential Applicants for any Private College using LightGBM. *2020 International Conference on Innovative Trends in Information Technology, ICITIIT 2020*. <https://doi.org/10.1109/ICITIIT49094.2020.9071525>
- Sujatha, R., Aarthy, S. L., Chatterjee, J. M., Alaboudi, A., & Jhanjhi, N. Z. (2021). A Machine Learning Way to Classify Autism Spectrum Disorder. *International Journal of Emerging Technologies in Learning*, *16*(6). <https://doi.org/10.3991/ijet.v16i06.19559>

- Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*, 7(1). <https://doi.org/10.1186/s40537-020-00349-y>
- Varshini, G. D., & Chinnaiyan, R. (2020). Optimized machine learning classification approaches for prediction of autism spectrum disorder. Dalam *Ann Autism Dev Disord*. 2020;
- Wang, J., Li, P., Ran, R., Che, Y., & Zhou, Y. (2018). A short-term photovoltaic power prediction model based on the Gradient Boost Decision Tree. *Applied Sciences (Switzerland)*, 8(5). <https://doi.org/10.3390/app8050689>
- Wang, R., Shi, Y., & Cao, W. (2019). GA-SURF: A new Speeded-Up robust feature extraction algorithm for multispectral images based on geometric algebra. *Pattern Recognition Letters*, 127. <https://doi.org/10.1016/j.patrec.2018.11.001>
- Xie, S., Karlsson, H., Dalman, C., Widman, L., Rai, D., Gardner, R. M., Magnusson, C., Schendel, D. E., Newschaffer, C. J., & Lee, B. K. (2019). Family History of Mental and Neurological Disorders and Risk of Autism. *JAMA network open*, 2(3), e190154. <https://doi.org/10.1001/jamanetworkopen.2019.0154>
- Zhang, D., & Gong, Y. (2020). The Comparison of LightGBM and XGBoost Coupling Factor Analysis and Prediagnosis of Acute Liver Failure. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3042848>
- Zhang, L., Wang, J., & An, Z. (2023). Vehicle recognition algorithm based on Haar-like features and improved Adaboost classifier. *Journal of Ambient Intelligence and Humanized Computing*, 14(2), 807–815. <https://doi.org/10.1007/s12652-021-03332-4>

Lampiran





UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax. (061) 7366998 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: univ_medanarea@uma.ac.id

Nomor : 428 /FT.6/01.10/VI/2023 6 Juni 2023
Lamp : -
Hal : Penelitian Dan Pengambilan Data Tugas Akhir

Yth. Pimpinan Bundaku Autism Clinic Center
Jln. Ismail Harun No.43/26 Tembung
Di
Deli Serdang

Dengan hormat,
Kami mohon kesediaan Bapak/Ibu berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	N A M A	N P M	PRODI
1	Yorris Siagian	198160014	Teknik Informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir pada perusahaan/Instansi yang Bapak/Ibu Pimpin.

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan ilmiah dan Skripsi yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul penelitian :

Klasifikasi Autis Menggunakan Ekstraksi Fitur *Speed Up Robust Feature* (SURF) dengan Algoritma *Boosting*

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.

Tembusan :
1. Ka. BAMAI
2. Mahasiswa
3. File


Dr. Ruddy S. Kom, M. Kom



BUNDABU AUTISM CLINIC CENTER

Jl. Ismail Harun No. 26/45 Titi Sewa Bandar Khalifah, Kec. Percut Sei Tuan
Kab. Deli Serdang Sumatera Utara. Kode Pos 20571 Contact Person : 0821 6185 4593

Melayani Anak-anak Berkebutuhan Khusus

- Autisme
- Terlambat Bicara
- Hyperaktif
- Kesulitan Belajar

Bandar Khalifah, 27 Juni 2023

Nomor : 038 / B-ACC/06/2023
Hal **Penelitian dan Pengambilan Data Tugas Akhir**
Kepada Yth,
Dekan
Fakultas Teknik
Universitas Medan Area

Tempat

Dengan hormat,
Berdasarkan surat nomor 428/FT.6/01.10/VI/2023 perihal Penelitian dan Pengambilan Data Tugas Akhir, maka dengan surat ini kami sampaikan bahwa mahasiswa dengan data sebagai berikut :

Nama : YORRIS SIAGIAN
NPM : 198160014
Program Studi : Teknik Informatika
Waktu pelaksanaan : 07 Juni 2023 s/d 27 Juni 2023

Telah menyelesaikan Riset atau Penelitian dan Pengambilan Data Tugas Akhir pada BUNDABU AUTISM CLINIC CENTER dengan tetap menjaga disiplin dan aset pribadi Center serta mentaati protokol kesehatan.

Demikian kami sampaikan, atas perhatian dan kerjasapannya kami ucapkan terimakasih

BUNDABU AUTISM CLINIC CENTER
PIMPINAN


SITI VIEN KUSALA DHEWI AZ, S.Pd



UNIVERSITAS MEDAN AREA

FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax. (061) 7366998 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: univ_medanarea@uma.ac.id

Nomor : 679/FT.6/01.10/IX/2023
Lamp : -
Hal : Perpanjang SK Pembimbing Tugas Akhir

20 September 2023

Yth. Pembimbing Tugas Akhir
Muhathir, ST, M. Kom
di
Tempat

Dengan hormat,
Sehubungan telah berakhirnya waktu masa berlaku SK pembimbing nomor 393/FT.6/01.10/V/2023 tertanggal 29 Mei 2023 maka perlu diterbitkan kembali SK Pembimbing Skripsi baru atas nama mahasiswa berikut :

N a m a : Yorris Siagian
N P M : 198160014
Jurusan : Teknik Informatika

Oleh karena itu kami mengharapkan kesediaan saudara :

Muhathir, ST, M. Kom (Sebagai Pembimbing)

Adapun Tugas Akhir Skripsi berjudul :

“Klasifikasi Autis Menggunakan Ekstraksi Fitur *Speed Up Robust Feature* (SURF) dengan Algoritma *Boosting*”

SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.

Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.

Dekan,

Dr. Rahmad Syah, S. Kom, M. Kom

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split, learning_curve
from sklearn.metrics import confusion_matrix, classification_report,
accuracy_score, precision_score, recall_score, f1_score, fbeta_score
from sklearn.ensemble import AdaBoostClassifier
import seaborn as sns

# Membaca data
df = pd.read_csv("dataset_features.csv")

# Memisahkan fitur dan label
X = df.drop('Label', axis=1).values
y = df['Label'].values

# Membagi data latih dan data uji
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# Melatih model AdaBoost
model = AdaBoostClassifier(n_estimators=50, learning_rate=0.1)
model.fit(X_train, y_train )

# Memprediksi data uji
y_pred = model.predict(X_test)

# Menghitung confusion matrix
cm = confusion_matrix(y_test, y_pred)
print("Confusion Matrix:")
print(cm)

# Menampilkan confusion matrix
plt.figure(figsize=(8, 5))
sns.heatmap(cm, annot=True, cmap="YlGnBu", fmt='d', cbar=False,
annot_kws={"size": 16})
plt.xticks([0.5, 1.5], ['Autis', 'Normal'], fontsize=12)
plt.yticks([0.5, 1.5], ['Autis', 'Normal'], fontsize=12)
plt.xlabel("Predicted Label", fontsize=14)
plt.ylabel("True Label", fontsize=14)
plt.show()

# Menampilkan classification report
print("Classification Report:")

```



```

print(classification_report(y_test, y_pred))

# Menghitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average=None)
recall = recall_score(y_test, y_pred, average=None)
f1 = f1_score(y_test, y_pred, average=None)
fbeta = fbeta_score(y_test, y_pred, beta=2.0, average=None)

print("Accuracy:", accuracy)
print("Precision - Autis:", precision[0])
print("Precision - Normal:", precision[1])
print("Recall - Autis:", recall[0])
print("Recall - Normal:", recall[1])
print("F1 Score - Autis:", f1[0])
print("F1 Score - Normal:", f1[1])
print("Fbeta Score - Autis:", fbeta[0])
print("Fbeta Score - Normal:", fbeta[1])

# Fungsi untuk plot learning curves
def plot_learning_curves(model, X_train, y_train, X_test, y_test):
    train_sizes, train_scores, test_scores = learning_curve(
        model, X_train, y_train, cv=5,
        scoring='accuracy',
        n_jobs=-1,
        train_sizes=np.linspace(0.01, 1.0, 50),
        shuffle=True,
        random_state=42
    )

    # menghitung mean dan deviasi standar untuk skor set pelatihan
    train_mean = np.mean(train_scores, axis=1)
    train_std = np.std(train_scores, axis=1)

    # menghitung mean dan deviasi standar untuk skor set pengujian
    test_mean = np.mean(test_scores, axis=1)
    test_std = np.std(test_scores, axis=1)

    # Plot kurva pembelajaran
    plt.figure(figsize=(10,6))
    plt.plot(train_sizes, train_mean, label='Training accuracy',
             color='blue', marker='o')
    plt.fill_between(train_sizes, train_mean - train_std, train_mean +
                    train_std, color='blue', alpha=0.15)

```

```
plt.plot(train_sizes, test_mean, label='Testing accuracy',
color='green', marker='s')
plt.fill_between(train_sizes, test_mean - test_std, test_mean +
test_std, color='green', alpha=0.15)

plt.title('adaboost')
plt.xlabel('Training Data Size')
plt.ylabel('Accuracy Score')
plt.legend(loc='lower right')
plt.grid()

# Memanggil fungsi dengan model AdaBoost
plot_learning_curves(model, X_train, y_train, X_test, y_test)
plt.show()
```

