

## PENINGKATAN PERFORMA ENSEMBLE LEARNING PADA SEGMENTASI SEMANTIK GAMBAR DENGAN TEKNIK OVERSAMPLING UNTUK CLASS IMBALANCE

Arie Nugroho<sup>\*1</sup>, M.Arief Soeleman<sup>2</sup>, Ricardus Anggi Pramunendar<sup>3</sup>, Affandy<sup>4</sup>, Aris Nurhindarto<sup>5</sup>

<sup>1,2,3,4,5</sup>Universitas Dian Nuswantoro, Semarang

Email: <sup>1</sup>arieunp81@gmail.com, <sup>2</sup>arief22208@gmail.com, <sup>3</sup>ricardus.anggi@dsn.dinus.ac.id,

<sup>4</sup>affandy@dsn.dinus.ac.id, <sup>5</sup>arisnurhindarto@dsn.dinus.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 20 Desember 2023, diterima untuk diterbitkan: 26 Juli 2023)

### Abstrak

Perkembangan teknologi dan gaya hidup manusia yang semakin tinggi menghasilkan data-data yang berlimpah. Data-data tersebut dapat berbentuk data yang terstruktur dan tidak terstruktur. Data gambar termasuk dalam data yang tidak terstruktur. Aktivitas dan objek yang terekam dalam suatu gambar beraneka ragam. Secara normal, mata manusia dapat dengan mudah membedakan antara *foreground* dan *background* dari suatu gambar, tetapi komputer membutuhkan pembelajaran dalam membedakan keduanya. Segmentasi gambar adalah salah satu bidang dalam *computer vision* yang membahas bagaimana cara komputer mempelajari dan mengenali segmen dari suatu gambar sesuai label yang ditentukan. Dalam kenyataannya banyak data yang mempunyai *class* atau label yang tidak seimbang, tentunya akan mempengaruhi tingkat akurasi dari suatu prediksi. Dalam riset ini membahas bagaimana meningkatkan akurasi segmentasi semantik gambar pada metode *ensemble learning* untuk menangani masalah data yang tidak seimbang dalam segmentasi gambar. Teknik yang digunakan adalah sintesis *oversampling* sehingga menghasilkan data yang seimbang dan akurasi yang tinggi. Metode *ensemble learning* yang digunakan adalah *Random Forest* dan *Light Gradient Boosting Machine* (LGBM). Dengan menggunakan dataset *Penn-Fudan Database for Pedestrian* yang mengandung *imbalanced class*. Penggunaan teknik sintesis *oversampling* dapat memperbaiki tingkat akurasi pada *class minoritas*. Pada algoritma *random forest* mengalami peningkatan akurasi sebesar 37 % sedangkan pada algoritma LGBM meningkat sebesar 41 %.

**Kata kunci:** *oversampling, class imbalanced, ensemble learning, segmentasi semantik gambar*

## IMPROVED PERFORMANCE OF ENSEMBLE LEARNING ON SEMANTIC SEGMENTATION OF IMAGES WITH OVERSAMPLING TECHNIQUES FOR CLASS IMBALANCE

### Abstract

The development of technology and the increasingly high lifestyle of humans produce abundant data. These data can be in the form of structured and unstructured data. Image data is included in unstructured data. The activities and objects recorded in a picture are varied. Normally, the human eye can easily distinguish between the foreground and background of an image, but computers need learning to distinguish between the two. Image segmentation is one of the fields in computer vision that discusses how computers learn and recognize segments of an image according to specified labels. In reality, a lot of data has unbalanced classes or labels, of course, it will affect the accuracy of a prediction. This research discusses how to improve the accuracy of image semantic segmentation in the ensemble learning method to deal with the problem of unbalanced data in image segmentation. The technique used is synthetic oversampling so as to produce balanced data and high accuracy. The ensemble learning methods used are Random Forest and Light Gradient Boosting Machine (LGBM). By using the Penn-Fudan Database for Pedestrian dataset which contains a imbalanced class. The use of synthetic oversampling techniques can improve the level of accuracy in minority classes. The random forest algorithm experienced an increase in accuracy by 37% while the LGBM algorithm increased by 41%.

**Keywords:** *oversampling, class imbalanced, ensemble learning, semantic image*

### 1 PENDAHULUAN

Seiring perkembangan teknologi, *Artificial Intelligence* (AI) berkembang menjadi beberapa

bidang ilmu yang lebih spesifik, di antaranya adalah *machine learning*. *Machine Learning* adalah bagian dalam AI yang bidang ilmunya membahas tentang

metode komputasi yang mampu meningkatkan kinerja dari pengalaman sebelumnya yang didapat dari proses *training* (Lu & Li, 2020). *Machine Learning* terdiri dari *supervised* dan *unsupervised*. *Supervised learning* membutuhkan input manual dalam data *training* sebagai label klasifikasinya, kemudian digunakan untuk membuat model yang akan dibandingkan dengan data *testing* untuk menghitung akurasi. *Unsupervised learning* merupakan kebalikan dari *supervised learning* dimana tidak memerlukan label. Dalam *Supervised learning* atau klasifikasi, data *training* digunakan untuk melakukan prediksi label yang bersifat nominal atau diskrit pada data *testing*. Bagian dalam *machine learning* dengan konsep jaringan saraf (*neural network*) yang mempunyai lapisan lebih dalam dikenal dengan *Deep Learning*. *Deep learning* digunakan untuk data gambar, video, suara dan dengan jumlah data yang besar. *Deep learning* digunakan karena semakin meningkatnya variasi data gambar di internet, belum diberi label dan semakin sulit untuk mengenali kategori yang sesuai atau benar, karena perbedaan skala, sudut pandang, rotasi dan *background* yang mengandung *noise* (Yang et al., 2021). Contoh penerapan *deep learning* adalah analisis gambar yang berkaitan dengan medik atau kesehatan, misalnya kasus klasifikasi gambar, deteksi dan segmentasi pada gambar retina, jantung, perut dan yang lain (Litjens et al., 2017).

Segmentasi gambar adalah proses yang bertujuan untuk mendapatkan objek terkandung di dalam suatu gambar atau membagi gambar menjadi beberapa wilayah atau area dengan setiap objek atau wilayah yang memiliki atribut yang sama (Syahrir et al., 2018). Pada segmentasi dengan data berupa gambar, AI membagi gambar menjadi beberapa area sesuai kedekatan jenis pixelnya, sehingga dapat memperkirakan bentuk dari suatu objek. Inputnya berupa gambar dan outputnya adalah gambar yang sudah tersegmentasi berdasarkan perkiraan objek dalam foreground dan backgroundnya. Dalam segmentasi gambar ada dua pendekatan, yaitu klasik dan modern. Pendekatan secara klasik, menggunakan metode *thresholding*, *region* dan *clustering*. Pendekatan secara modern menggunakan metode semantik dan *instance*. Segmentasi gambar dengan metode semantik merupakan proses segmentasi gambar dengan class yang sudah disediakan, jadi proses segmentasinya dapat membedakan atau klasifikasi pixel sesuai dengan class atau maknanya, contohnya adalah jika dalam sebuah gambar pemandangan alam akan dilakukan klasifikasi pixel sehingga dapat diketahui segmentasi air, daratan, pohon atau yang lainnya. Tujuannya adalah memberi label dari setiap pixel sesuai dengan class yang telah ditentukan sebelumnya. Penggunaan segmentasi semantik antara lain untuk diagnosa medis, membantu atau memandu disabilitas, identifikasi area dalam suatu

kondisi alam dan yang lainnya. Di dalam segmentasi gambar secara semantik tidak dapat menentukan berapa banyak class dalam segmentasinya. Untuk dapat menentukan banyaknya object menggunakan segmentasi *instance*. Dalam segmentasi ini tidak hanya membedakan area berdasarkan labelnya, tetapi dapat menentukan banyaknya object.

Pada penelitian ini membahas tentang segmentasi semantik gambar untuk membedakan area sesuai labelnya. Dalam penelitian ini lebih khusus membahas tentang segmentasi semantik yang berkaitan dengan class imbalanced atau class dari pixel yang tidak seimbang pada class yang diberikan dalam suatu gambar. Imbalanced class dalam segmentasi semantik gambar akan terjadi jika pixel yang sesuai dengan class tertentu jauh lebih banyak (mayoritas) daripada pixel dari satu atau lebih class (minoritas), kelangkaan class minoritas dalam data pelatihan menghambat pembelajaran dan pelabelan yang akurat, karena model yang dipelajari akan cenderung mengklasifikasikan sebagian besar pixel sebagai anggota class mayoritas, hal ini mengakibatkan kurang tepatnya hasil segmentasi (Small & Ventura, 2017). Class Imbalanced dapat terjadi karena beberapa kemungkinan, yaitu kesalahan atau kekurangan dalam pengambilan data, kemungkinan yang lain adalah dari awal datanya memang sudah tidak seimbang (imbalanced). Pada kasus tidak seimbangnya class, banyak cara yang bisa digunakan untuk mengatasinya, antara lain memberikan pembobotan (*weight*) pada kelas minoritas, teknik *sampling*, penentuan *similarity loss function* dan *focal lost*. Kombinasi penggunaan teknik *deep learning* dan pemilihan *lost function* telah dibuktikan dapat mengatasi tidak seimbangnya data, namun memerlukan waktu dan hardware yang lebih untuk memprosesnya (Hashemi et al., 2019). Cara lain untuk mengatasi kelas yang tidak seimbang adalah menggunakan *ensemble learning* (Liu & Wu, 2019). *Ensemble learning* berdasarkan *decision tree* merupakan gabungan dari beberapa model algoritma *machine learning* dalam bentuk pohon keputusan untuk menyelesaikan masalah class imbalanced pada klasifikasi, namun ketika suatu model yang dihasilkan belum mampu memberikan hasil yang akurat maka diperlukan suatu cara untuk membantu meningkatkan akurasi, yaitu dengan cara *hybrid* atau memadukan *ensemble learning* dengan cara lain untuk meningkatkan akurasi pada class imbalanced (Ruangthong & Jaiyen, 2016). Selain *ensemble learning*, cara untuk menyelesaikan masalah ketidakseimbangan kelas pada klasifikasi, adalah melakukan penyeimbangan data-level atau tingkat data dengan melakukan *pre-processing* pada dataset, yaitu dengan melakukan teknik *sampling* dengan sintesis *oversampling* (Guo et al., 2019). Dalam teknik ini, dataset dengan data yang memiliki jumlah kelas yang berbeda (minoritas) akan disamakan jumlahnya dengan kelas mayoritas,

sehingga antar kelas akan memiliki jumlah data yang sama.

Berdasarkan uraian di atas, maka akan dilakukan upaya untuk meningkatkan akurasi ensemble learning untuk menangani class imbalanced pada segmentasi semantik gambar, yaitu dengan menggabungkannya dengan tahap pre-processing, yaitu dengan menyeimbangkan ketidakseimbangan data sebelum diolah menggunakan ensemble learning (EL). Dalam penelitian ini menggunakan 2 (dua) algoritma EL, yaitu Random Forest dan Light Gradient Boosting Machine (LGBM) serta menggunakan teknik sampling dengan sintesis oversampling untuk menyeimbangkan jumlah piksel antara class minoritas dan mayoritas pada data gambar yang akan diolah, sehingga hasil yang didapatkan lebih baik dari penelitian sebelumnya ditunjukkan dengan tingkat akurasi yang lebih tinggi. Tujuan dari penelitian ini adalah memperoleh akurasi yang lebih baik dari metode EL untuk menangani masalah class imbalanced dengan menggunakan teknik sintesis oversampling dalam segmentasi semantik pada gambar. Dalam penelitian ini hanya membahas tentang cara meningkatkan akurasi class minoritas pada segmentasi semantik pada gambar pedestrian dengan 1 objek pejalan kaki dan background. Berdasarkan latar belakang yang telah diuraikan, yaitu bagaimana meningkatkan akurasi metode EL pada segmentasi semantik gambar sebagai salah satu cara menangani class imbalanced, rumusan masalahnya adalah apakah penggunaan teknik sintesis oversampling pada segmentasi semantik gambar pedestrian dapat meningkatkan akurasi metode EL untuk mengatasi masalah class imbalanced.

## 2 METODE PENELITIAN

### 2.1 Data Penelitian

Dataset yang digunakan dalam penelitian ini adalah dataset publik, yaitu gambar pedestrian atau pejalan kaki. Dataset tersebut mempunyai 2 (dua) class, yaitu class pejalan kaki dan background (objek selain pejalan kaki). Objek dalam dataset memiliki 1 dan beberapa objek pejalan kaki. Background dari dataset ini terdiri dari banyak objek, yaitu mobil, pohon dan yang lain. Dalam dataset ini ada beberapa citra pedestrian dengan objek seorang atau beberapa pejalan kaki dan backgroundnya. Fokus dari dataset ini hanya pada segmentasi objek dari pejalan kaki. Dalam proses segmentasi semantik dibutuhkan citra groundtruth untuk mengetahui berapa dan apa saja label yang akan dijadikan acuan. Pada citra-citra tersebut ada yang terdiri dari single objek (pejalan kaki) dan juga ada yang lebih. Pada setiap objek pejalan kaki mempunyai warna masking yang berbeda-beda. Pada segmentasi semantik seharusnya setiap objek dengan kelas yang sama mempunyai warna yang sama, sehingga dalam penelitian ini

hanya fokus pada citra dengan objek tunggal dari pejalan kaki karena merupakan ciri dari segmentasi semantik yang citra groundtruthnya 1 kelas hanya mempunyai 1 warna.

### 2.2 Metode yang diusulkan

Oskar Taubert, dkk memperkenalkan pendekatan baru untuk mengatasi masalah ketidakseimbangan kelas yang sangat relevan dengan masalah segmentasi gambar. Dalam contoh yang dieksplorasi dalam loss scheduling dalam kasus ini secara konsisten mampu mendorong model dalam optimasi segmentasi gambar, yang lebih menekankan pada kelas yang kurang terwakili (minority class). Dataset yang digunakan untuk pelatihan adalah gambar Cityscapes, dimana model high recall cocok untuk masalah di mana false negative di kelas yang kurang terwakili memiliki nilai yang tinggi, misalnya dalam deteksi patologi biomedis.. Pengukuran atau metrics dalam riset ini menggunakan confusion matriks dengan akurasi per-kelas dan F1-Score dan recall. Hasil dari riset ini adalah loss berdasarkan reweighting cross entropy untuk fokus pada kasus yang sulit dapat menghasilkan model dengan kinerja yang lebih baik, terutama dengan melakukan pengaturan hyper-parameter-nya. Namun re-weighting tampaknya berdampak buruk pada kalibrasi model terlatih. Meskipun loss scheduling tidak dapat menghasilkan kinerja model yang telah dilatih sebelumnya pada himpunan data yang lebih besar, harapan dalam riset ini adalah dapat dicapainya berbagai pilihan loss function atau loss scheduling yang lebih rumit (misalnya berosilasi) setidaknya untuk beberapa himpunan data. Loss scheduling juga dapat meningkatkan kinerja pada masalah seperti pelabelan tepi grafik atau prediksi peta fitur yang padat (Taubert et al., 2020).

Xiaoling Xia dkk membahas tentang bagaimana cara untuk menyelesaikan masalah dataset yang tidak seimbang dalam segmentasi semantik pada gambar. Riset ini mengambil kumpulan data yang disediakan dalam tantangan deteksi kapal airbus di Kaggle sebagai contoh untuk mengeksplorasi metode yang mudah dan efektif untuk segmentasi pada kumpulan gambar dengan ketidakseimbangan kelas. Metode yang digunakan adalah U-Net dengan model ResNets yang telah dilatih sebelumnya, dan mencoba berbagai metode untuk mengeksplorasi fitur set. Dalam proses pelatihan ResNets, riset ini mengusulkan struktur blok konvolusional baru yang terinspirasi oleh urutan Fibonacci. Pengukuran yang digunakan dalam riset ini adalah dice coefficient dan IoU. Pada akhirnya, nilai mF2 dari model yang dilatih dalam riset ini mencapai hasil yang baik, yang lebih baik daripada model pelatihan gabungan ResNets dan U-Net34 biasa. Selain itu, parameter pelatihan kurang dari itu. Riset ini meyakini bahwa metode pelatihan yang sederhana dan efektif ini akan membawa

kenyamanan bagi para peneliti di bidang terkait dengan metode yang diusulkan block konvolusi berdasarkan Fibonacci dan menghasilkan metode segmentasi yang efektif dengan perbedaan yang signifikan antara gambar background dan foregroundnya (Xia et al., 2019).

Sajjad Shumaly, dkk membahas tentang ketidak seimbangan kelas menggunakan teknik sampling dengan teknik boosting dan bagging dimana merupakan algoritma ensemble learning pada klasifikasi atau prediksi pada data customer churn. Pada penelitian ini, masalahnya terkait dengan klasifikasi pelanggan menjadi dua kelompok, pelanggan setia dan pelanggan yang berniat untuk berhenti menggunakan layanan organisasi. Awalnya, lima algoritma pembelajaran mesin digunakan dalam situasi di mana datanya tidak seimbang, dan dalam hal ini, hasilnya tidak memuaskan. Algoritma yang digunakan adalah decision tree, support vector machine, multilayer perceptron, gradien boosting dan random forest. Teknik sampling yang digunakan adalah undersampling dan oversampling, yaitu dengan mengurangi kelas mayoritas sehingga dapat seimbang dengan kelas minoritas dan juga menambahkan data pada area minoritas sehingga sama dengan banyaknya data pada kelas mayoritas. Dalam penelitian ini, karena ketidakseimbangan data, akurasi bukan indikator yang baik untuk mengukur kinerja algoritma. Oleh karena itu, penting untuk menggunakan indikator yang sensitif terhadap simetri data. Untuk tujuan ini, indeks AUC akan digunakan, yang menunjukkan area di bawah diagram dalam kurva ROC. Kurva ROC adalah metode untuk menilai kinerja pengklasifikasi dari dua klasifikasi. Dalam kurva ROC sumbu vertikal adalah TPR, yaitu kadang-kadang disebut sensitivitas dan mewakili tingkat positif yang sebenarnya. Juga, sumbu horizontal adalah FPR yang menunjukkan tingkat false positif. Hasil dari riset ini adalah performa random forest dan gradien boosting memperoleh akurasi tertinggi di antara algoritma klasifikasi yang lain (Shumaly et al., 2020).

Metode yang diusulkan pada penelitian ini adalah menerapkan metode transfer learning yang menggabungkan VGG-16 dan algoritma EL serta teknik sintetis oversampling. VGG-16 digunakan sebagai alat untuk ekstraksi fitur yang kemudian proses klasifikasi pikselnya menggunakan algoritma dari metode EL. Teknik sintetis oversampling digunakan untuk menyeimbangkan antara class minoritas dan mayoritas. Sintetis oversampling hanya dikenakan pada class minoritas dalam data training, sehingga data training dari class minoritas dan mayoritas dapat seimbang dan diharapkan dapat menangani masalah class Imbalanced pada segmentasi gambar.

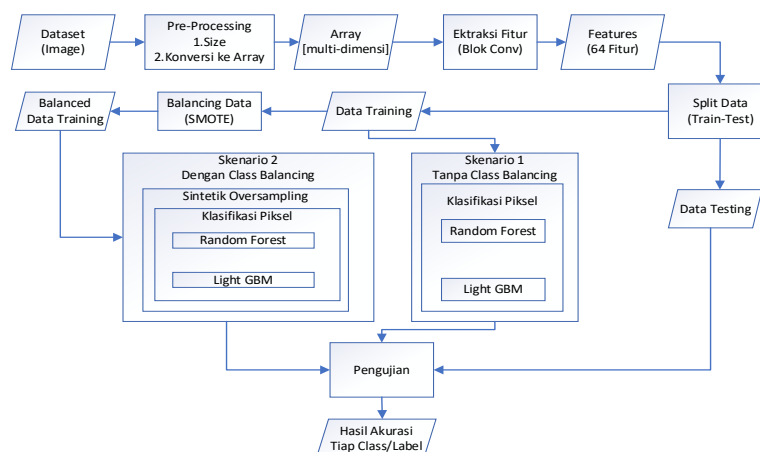
Alasan menggunakan VGG-16 sebagai alat untuk ekstraksi fitur karena arsitektur ini meraih akurasi yang tinggi pada dataset imageNet yang

memiliki lebih dari 14 juta gambar dengan 1000 class. Arsitektur ini telah dilatih selama minggu menggunakan GPU NVIDIA Titan Black. Alasan menggunakan metode EL karena dalam metode ini mempunyai algoritma-algoritma tersebut termasuk pembelajaran secara gabungan yang cocok digunakan untuk menyelesaikan kasus imbalanced class (He & Ma, 2013). Alasan menggunakan metode oversampling ini adalah dalam proses klasifikasi dengan machine learning akan lebih baik jika menggunakan data training yang sebanyak mungkin dan memiliki persebaran data yang normal sehingga model dari machine learning yang telah dihasilkan akan mempunyai bahan atau data yang cukup untuk proses belajarnya kemudian jika model tersebut digunakan dalam proses validasi dengan data testing, maka akan menghasilkan akurasi yang lebih baik dibandingkan dengan model yang dibentuk dari data yang kurang. Sedangkan jika menggunakan teknik undersampling maka akan kehilangan beberapa informasi atau data yang mungkin penting untuk proses pembelajaran. Teknik oversampling menggunakan sintetis karena untuk mengurangi resiko overfitting dibandingkan menggunakan teknik random yang melakukan duplikasi data acak secara langsung dan mempunyai resiko overfitting.

Pada tahap ekstraksi fitur menggunakan 2 blok layer konvolusi dari VGG-16 yang akan menghasilkan fitur-fitur sesuai dengan arsitektur dari 2 blok layer konvolusi. Dari fitur-fitur yang telah dihasilkan akan displit atau dipecah menjadi data training sebanyak 80 % dan testing sebanyak 20 %. Pada tahap klasifikasi, menggunakan dataset yang belum dan yang telah dimodifikasi dengan teknik sintetis piksel untuk menyeimbangkan class minoritas dengan mayoritas, kemudian proses klasifikasi datanya menggunakan beberapa algoritma dalam metode EL, yaitu Random Forest dan LGBM.

Pada gambar 1 langkah pertama adalah melakukan pre-processing terhadap dataset gambar, yaitu gambar train dan ground truth (GT). Pemrosesan awal yang dilakukan adalah penyamaan ukuran dari data gambar yang digunakan, kemudian mengubah image menjadi angka dalam bentuk array 2 dimensi. Langkah berikutnya adalah menentukan  $X_{train}$  dari array gambar training dan  $y_{train}$  dari array gambar mask (GT).

Langkah berikutnya adalah tahap ekstraksi fitur yang diproses dengan blok convolutional dari VGG-16 dengan weights dari imageNet. Dalam langkah ini melakukan prediksi model dari data train, yang akan menghasilkan 64 fitur. Berikutnya adalah menentukan X dan Y dari fitur yang telah dihasilkan dari proses ekstraksi fitur. Berikutnya adalah menampilkan fitur-fitur unik dari piksel-piksel gambar untuk dijadikan class. Dari langkah tersebut dihasilkan 2 (dua) fitur unik, yaitu area dengan warna hitam diwakili dengan angka 0 dan area putih diwakili dengan angka 1.



Gambar 1. Metode yang diusulkan

Dalam langkah ini juga ditampilkan banyaknya piksel dari setiap fitur unik yang telah dihasilkan, sehingga dapat diketahui mana area atau class yang mayoritas dan mana yang minoritas. Berikutnya adalah membagi data menjadi training dan testing.

Alur pada penelitian ini dibagi menjadi 2 skenario, yaitu tanpa dan dengan class balancing. Skenario pertama adalah tanpa class balancing, yaitu klasifikasi piksel langsung dari data training aslinya. Klasifikasi piksel dengan metode EL dengan algoritma Random Forest dan LGBM, kemudian dilakukan validasi dengan data testing. Dari skenario pertama akan diketahui class mana yang mempunyai akurasi yang paling rendah. Dari hasil penelitian, class mayoritas lebih mudah untuk dipelajari dalam pembuatan model machine learning dibuktikan dengan akurasi yang tinggi, tapi untuk class minoritas akan lebih sulit untuk dipelajari karena jumlah yang lebih sedikit dibuktikan dengan akurasi yang lebih rendah, sehingga pada skenario kedua dilakukan proses class balancing data dengan melakukan penambahan data atau piksel pada class minoritas tersebut (oversampling), yaitu dengan sintesis oversampling (SMOTE).

Dalam gambar pedestrian, class pedestrian / pejalan kaki merupakan class minoritas, sedangkan class background adalah class mayoritas. Skenario kedua adalah melakukan class balancing dengan sintesis oversampling adalah dengan menyeimbangkan jumlah piksel dari class minoritas dengan membuat data baru berdasarkan kedekatan dari data-data yang sebelumnya. Sintesis oversampling membutuhkan waktu yang lebih lama daripada random, karena dalam sintesis dilakukan proses perhitungan untuk menentukan data-data baru. Teknik sintesis dilakukan untuk mengurangi masalah overfitting. Setelah melakukan langkah oversampling, akan diproses dengan metode EL menggunakan 2 algoritma, kemudian melakukan validasi dengan data testing. Langkah terakhir adalah melakukan pengujian model dengan data yang baru atau data yang belum pernah digunakan sebagai training dan testing kemudian menampilkan hasil akurasi dari setiap label atau class.

## 2.3 Pre-processing

Pemrosesan awal yang dilakukan pada tahap ini adalah mengatur ukuran gambar yang akan diproses, yaitu standarisasi ukuran tinggi (*height*) dan lebar (*width*). Pada penelitian ini menggunakan pengaturan standar untuk gambar input dengan ukuran 300x300. Setelah menentukan ukuran standar dari gambar, berikutnya adalah memasukkan gambar training ke dalam list gambar, kemudian melakukan konversi list gambar ke dalam bentuk numpy array. Hasilnya adalah variable `train_images` dengan tipe array `uint8` dengan ukuran (jumlah\_image, ukuranWidth, ukuranHeight, Channel). Isi dari variabel `train_images` berisi angka 0 s/d 255 dalam bentuk array multi dimensi. Hal yang sama akan dilakukan untuk gambar mask (GT), yaitu memasukkan citra GT ke dalam list, kemudian mengkonversikan ke dalam bentuk numpy array di mana menghasilkan variable `train_mask` dengan tipe data array `uint8` dengan ukuran (jumlah\_image, lebar, tinggi). Isi dari variabel `train_mask` adalah angka 0 dan 1 dengan keterangan 0 berarti background dan 1 menyatakan objek pejalan kaki.

## 2.4 Ekstraksi Fitur

Berikutnya adalah tahap ekstraksi fitur dari data latih dan GT, yang dalam penelitian ini menggunakan salah satu arsitektur Convolutional Neural Network (CNN), yaitu VGG 16 dengan pre-trained weight imageNet dengan tipe konvolusi untuk citra 2 dimensi. Pada tahap ini model VGG-16 yang digunakan untuk tahap ekstraksi fitur hanya pada proses konvolusinya saja tanpa layer dense atau tanpa layer classifernya. Berikut adalah potongan coding dari penggunaan VGG-16 tanpa layer klasifikasinya dan load weight dari imageNet untuk proses generate dalam ekstraksi fitur. Model convolutional VGG-16 yang terdiri dari input layer dan 5 blok konvolusi. Pada 2-layer konvolusi yang pertama memiliki 64 filter. Pada 2 layer berikutnya, yaitu 3 dan 4 mempunyai 128 filter. Pada layer berikutnya, yaitu 4,5 dan 6 memiliki 256 filter. Pada sisa layer berikutnya memiliki 512 filter. Dalam

proses ekstraksi fitur dalam penelitian ini hanya menggunakan 2-layer konvolusi pertama, yaitu dengan 64 filter. Pada layer dense atau layer klasifikasi akan digantikan dengan metode ensemble learning (EL) untuk proses klasifikasi pikselnya. Gabungan dari arsitektur CNN sebagai feature extractor dan algoritma lain sebagai alat klasifikasinya disebut transfer learning. Hasil dari proses ekstraksi fitur adalah array multi dimensi yang berisi bilangan float/decimal dengan jumlah kolom 64 fitur / filter. Fitur yang berjumlah 64 berdasarkan dari arsitektur VGG-16 dimana pada blok conv 1 dan 2 mempunyai memiliki 64 fitur.

Berikutnya adalah proses generate fitur berdasarkan model baru yang telah dibuat dengan 2-layer blok konvolusi dari data  $X_{train}$  (data training) yang akan menghasilkan features dengan 64 fitur dengan tipe data Array multi-dimensi angka decimal. Proses berikutnya adalah melakukan reshape, yaitu mengubah array multi-dimensi menjadi array 2 dimensi yaitu bentuk baris dan kolom, hal ini dilakukan agar bisa diproses dengan algoritma klasifikasi. Hasil dari proses reshape features untuk variabel  $X$ , dimana mempunyai bentuk 2 dimensi dengan 64 fitur / kolom dengan tipe bilangan decimal dan untuk isi dari variabel  $Y$  didapat dari reshape  $y_{train}$  yang juga terdiri dari 1 kolom dengan tipe bilangan integer. Bentuk Perubahan bentuk dari multi-dimensi menjadi 2 dimensi mengakibatkan bertambahnya banyaknya baris/row, yang didapat dari jumlah gambar train dikalikan dengan lebar dan panjang ukuran gambar. Tahap berikutnya adalah memasukkan  $X$  dan  $Y$  ke dataframe untuk proses klasifikasi dengan algoritma EL. Dari dataframe dapat diketahui ada berapa label atau class yang ada pada gambar GT dan banyak data dari setiap labelnya. Hasil dari proses tersebut adalah yaitu background dengan nilai 0 dan pejalan kaki dengan nilai 1. Kelas background memiliki jumlah data yang lebih banyak (mayoritas) terhadap kelas pejalan kaki (minoritas).

## 2.5 Klasifikasi dengan Metode Ensemble Learning

Metode ensemble learning bertujuan untuk meningkatkan kinerja suatu algoritma klasifikasi dengan menggabungkan beberapa classifier sebagai classifier baru dari suatu data kemudian menggabungkan beberapa prediksi untuk dipilih sebagai prediksi akhir, yang mana hasilnya akan lebih baik dari classifier awal, dimana dalam metode ini terdiri dari bootstrap aggregating (bagging), boosting dan stacking (Fernández et al., 2018). Salah satu contoh dari pendekatan bagging adalah algoritma random forest yang terdiri dari beberapa pohon keputusan (Nugroho et al., 2021). Bootstrapped dataset diambil dari dataset yang diambil secara acak yang kemudian setiap bootstrapped-nya diproses dengan masing-masing satu decision tree.

Setiap decision tree akan menghasilkan satu prediksi yang kemudian akan divotting untuk suara terbanyak. Hasil prediksi yang mempunyai suara terbanyak akan dijadikan sebagai prediksi akhir.

Boosting adalah salah satu pendekatan dari ensemble learning yang juga terdiri dari beberapa decision tree dimana mempunyai sifat berurutan dan pada setiap model yang dihasilkan tergantung dari model sebelumnya sehingga pada setiap model akan ada keterkaitan yang kemudian di-ensemble untuk menghasilkan prediksi terbaik. Algoritma yang termasuk dalam boosting adalah XGboost dan LGBM.

Light Gradient Boosting Machine (LGBM) adalah pengembangan juga dari decision tree yang dikembangkan untuk kecepatan memiliki yang lebih dalam proses training dan penggunaan memory yang kecil tapi masih dapat meraih akurasi yang tinggi. LGBM menumbuhkan daun pohon dengan membelah simpul daun menggunakan algoritma berbasis histogram, yang menghasilkan keuntungan besar, yaitu pada efisiensi dan konsumsi memori. Pertumbuhan pohon berdasarkan daun yang mengarah pada kompleksitas model yang lebih banyak, LGBM dapat menghasilkan perolehan akurasi yang lebih tinggi di setiap iterasinya (Liang et al., 2022).

Metode ensemble learning yang akan digunakan pada tahap klasifikasi ini adalah random forest dan LGBM. Kedua algoritma tersebut adalah pengembangan dari decision tree. Random forest dengan jumlah pohon sebanyak 50 dan 80 data training serta 20 % data testing. Langkah berikutnya adalah fit model yang membutuhkan waktu cukup lama sesuai dengan ukuran gambar dan banyaknya pohon estimatornya. Setelah proses fit model, berikutnya adalah prediksi data testing dengan model dari random forest yang telah dibuat. LGBM dengan parameter learning rate = 0.05, boosting type = dart, objective = multiclass, num\_leaves = 100, max\_depth = 10, num\_class = 2 dan 50 iterasi. Data train sebanyak 80 % dan sisanya untuk data testing.

## 3 HASIL DAN PEMBAHASAN

### 3.1 Skenario 1 (Tanpa Class Balancing)

Berdasarkan hasil klasifikasi dari 2 algoritma pada tahap sebelumnya dan masih menggunakan data yang belum seimbang, dihasilkan nilai *Intersection over Union* (IoU) dari setiap kelas yang tidak seimbang. Nilai IoU dari class background untuk kedua algoritma memperoleh hasil yang tinggi, sedangkan untuk class pejalan kaki masih mempunyai akurasi yang rendah. Akurasi setiap kelas didapat dari perbandingan nilai *values true positif* dan *true negative* dengan semua nilai dari *values*. Hasil dari perhitungan akurasi pada setiap kelas (IoU) pada dataset yang belum diseimbangkan ditunjukkan pada tabel 1.

Tabel 1. Perbandingan akurasi setiap class pada imbalanced dataset

algoritma	class	Nilai IOU (%)
1	BG	99
	PK	68
2	BG	99
	PK	44

Berdasarkan tabel 1 nilai IoU dari class background (BG) pada algoritma random forest (algoritma 1) dan LGBM (algoritma 2) menghasilkan nilai 99 %, hal ini dikarenakan class BG adalah kelas mayoritas. Nilai IoU dari class pejalan kaki (PK) pada algoritma 1 menghasilkan nilai 68 %, sedangkan untuk algoritma 2 mendapatkan nilai sebesar 44 %. Class PK mempunyai nilai IoU yang rendah karena merupakan kelas minoritas.

### 3.2 Skenario 2 (Dengan Class Balancing)

Pada skenario yang kedua, dilakukan penyeimbangan jumlah data pada kedua class ( BG dan PK) terhadap dataset yang telah diproses dalam ekstraksi fitur. Penyeimbangan data menggunakan metode oversampling, yaitu menambahkan data pada class minoritas, yaitu pada class PK sehingga mempunyai jumlah yang sama dengan class mayoritas, yaitu class BG. Metode oversampling yang digunakan adalah sintesis atau biasa disebut dengan *Synthetic Minority Over-sampling Technique* (SMOTE). Pada tahap ini jumlah data pada piksel dari semua class akan disamakan dengan cara sintesis. Hasil dari proses ini adalah setiap class pada label gambar mempunyai jumlah piksel atau data yang sama, sehingga dataset menjadi seimbang dan akan digunakan lagi dalam proses klasifikasi piksel dengan metode ensemble learning. Langkah pertama adalah mendapatkan data yang menjadi kelas minoritas, kemudian memisahkan antara data dan labelnya. Berikutnya adalah memanggil library imblearn untuk oversampling dengan SMOTE, kemudian melakukan fit resample dari data dan labelnya. Hasil dari perhitungan akurasi pada setiap kelas (IoU) pada dataset yang sudah diseimbangkan ditunjukkan pada tabel 2.

Tabel 2. Perbandingan akurasi setiap class pada balanced dataset

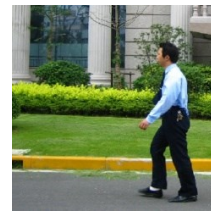
algoritma	class	Nilai IOU (%)
1	BG	99
	PK	99
2	BG	87
	PK	91

Dari tabel 2 dapat dijelaskan bahwa nilai IoU pada setiap kelas minoritas (PK) di semua algoritma mengalami peningkatan. Pada algoritma 1 class PK mengalami peningkatan nilai IoU yang semula hanya 68 % menjadi 99 %, dimana mengalami peningkatan sebesar 31 %, sedangkan pada algoritma 2 class PK juga mengalami peningkatan sebesar 47 %, yang semula hanya 44 % menjadi 91

%. Pada class BG di algoritma 2 mengalami penurunan nilai IoU sebesar 12 %.

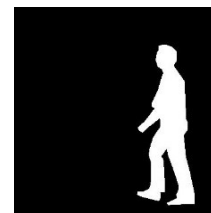
## 4 PENGUJIAN

Pada tahap ini akan ditampilkan hasil pengujian dengan data baru dari model yang telah dibuat tanpa dan dengan class balancing. Tahap pengujian dilakukan untuk menguji metode yang diusulkan dalam menangani data baru, dimana data tersebut belum pernah digunakan dalam proses training dan validasi. Berikut adalah gambar yang dijadikan data pengujian ditunjukkan pada gambar 3 dan groundtruthnya pada gambar 4.



Gambar 2. Data Pengujian

Pada gambar 2 ditunjukkan gambar pejalan kaki dengan *single object* (seorang pejalan kaki) dengan background bangunan, rumput dan tumbuhan lain.



Gambar 3. Data Groundtruth

Pada gambar 3 ditampilkan gambar groundtruth dari data pengujian. Fokus dari gambar groundtruthnya hanya pada objek pejalan kaki yang diwakili dengan warna putih, sedangkan background diwakili dengan warna hitam.

Tahap awal untuk pengujian adalah membaca gambar yang akan digunakan untuk testing model. Ukuran gambar testing juga harus disesuaikan (*resize*) sesuai ukuran pada data training. Berikutnya adalah melakukan prediksi gambar testing dengan model baru dari VGG-16 untuk tahap ekstraksi fitur diletakkan dalam variabel `X_test_feature`. Berikutnya adalah melakukan *reshape* ukuran menjadi 2 dimensi, yaitu menjadi data tabular (baris dan kolom). Hasil prediksi model dengan hasil dari ekstraksi fitur gambar yang dijadikan data testing pada tahap sebelumnya, kemudian dikembalikan ke bentuk gambar sebagai hasil segmentasi.

### 4.1 Pengujian Skenario 1 (tanpa class Balancing)

Pengujian skenario pertama menggunakan model dengan data yang tidak seimbang dengan



membandingkan 2 algoritma dan akan ditampilkan nilai IoU dari setiap kelasnya. Perbandingan akurasi setiap kelas ditunjukkan pada tabel 3.

Tabel 3. Perbandingan akurasi setiap class tahap pengujian pada imbalanced dataset

algoritma	class	Nilai IOU (%)
1	BG	94
	PK	53
2	BG	97
	PK	57

Pada tabel 3 ditampilkan perbandingan hasil akurasi pada setiap kelas ditunjukkan dengan nilai IoUnya untuk kedua algoritma dengan data yang tidak seimbang. Pada algoritma *random forest* objek pejalan kaki tidak tersegmentasi dengan sempurna, selain itu juga ditemukan *noise* pada *background*. Untuk nilai akurasi pada setiap kelas pada *background* mencapai akurasi sebesar 94 %, tapi pada kelas pejalan kaki hanya mencapai akurasi sebesar 53 %. Pada algoritma LGBM objek pejalan kaki juga kurang sempurna, pada bagian kepala nyaris tidak tersegmentasi dengan baik, tapi *noise* pada kelas *background* lebih sedikit dibandingkan *random forest*. Nilai akurasi pada setiap kelas pada algoritma LGBM mencapai akurasi sebesar 97 % untuk kelas *background* dan 57 % untuk kelas pejalan kaki. Hasil segmentasi di tahap pengujian dengan kedua algoritma pada data yang tidak seimbang (skenario 1) ditunjukkan pada tabel 4.

Tabel 4. Hasil segmentasi tahap pengujian pada imbalanced data

algoritma	hasil segmentasi
1	
2	

Pada tabel 4 ditunjukkan hasil segmentasi dengan data yang belum seimbang menggunakan 2 algoritma. Pada algoritma 1, yaitu *random forest* ditunjukkan objek pejalan kaki yang tersegmentasi dengan kurang sempurna dan masih ditemukan *noise* yang pada kelas *background*. Pada algoritma 2, yaitu LGBM objek pejalan kaki juga belum sempurna, pada bagian kepala nyaris tidak tersegmentasi dengan baik, *noise* pada kelas *background* juga ditemukan tapi tidak sebanyak jika dibandingkan dengan algoritma 1.

## 4.2 Pengujian Skenario 2 (dengan class balancing)



Pengujian skenario kedua menggunakan model dengan data yang telah diseimbangkan menggunakan metode sintesis oversampling dengan SMOTE. Dengan data yang seimbang serta menggunakan 2 algoritma akan ditampilkan nilai IoU dari setiap kelas ditunjukkan pada tabel 5.

Tabel 5. Perbandingan akurasi setiap class tahap pengujian pada balanced dataset

algoritma	class	Nilai IOU (%)
1	BG	94
	PK	76
2	BG	89
	PK	88

Pada tabel 5 ditunjukkan perbandingan hasil nilai IoUnya untuk kedua algoritma menggunakan data yang seimbang. Pada algoritma *random forest* objek pejalan kaki tidak tersegmentasi dengan sempurna, selain itu juga ditemukan *noise* pada *background*. Untuk nilai akurasi pada setiap kelas pada *background* mencapai akurasi sebesar 94 %, pada kelas pejalan kaki mencapai akurasi sebesar 76 %, akurasi ini lebih tinggi dibandingkan dengan skenario pertama. Hasil pengujian dengan algoritma LGBM pada kelas BG mengalami penurunan akurasi menjadi 89 %, sedangkan untuk kelas PK mengalami kenaikan menjadi 88 % dibandingkan dengan skenario pertama. Hasil dari perbandingan segmentasi dari gambar uji dengan data yang telah seimbang akan ditunjukkan pada tabel 6.

Tabel 6. Hasil segmentasi tahap pengujian pada balanced data

algoritma	hasil segmentasi
1	
2	

Pada tabel 6 ditunjukkan hasil segmentasi dengan data yang seimbang menggunakan 2 algoritma. Pada algoritma 1, yaitu *random forest* ditunjukkan objek pejalan kaki yang tersegmentasi dengan lebih sempurna dibandingkan dengan model *imbalanced data*, selain itu juga masih ditemukan *noise* yang lebih banyak pada kelas *background*. Pada algoritma LGBM objek pejalan kaki juga terlihat lebih sempurna dari model sebelumnya, pada bagian kepala nyaris tersegmentasi dengan baik, tapi *noise* pada kelas

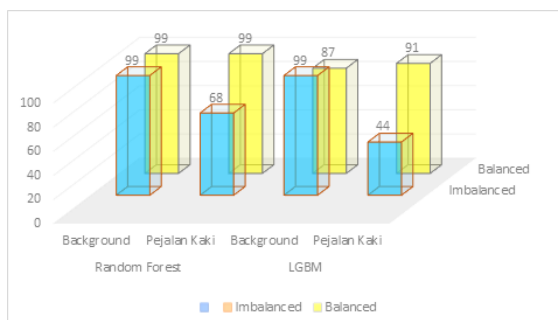


*background* lebih banyak dibandingkan dengan model sebelumnya.

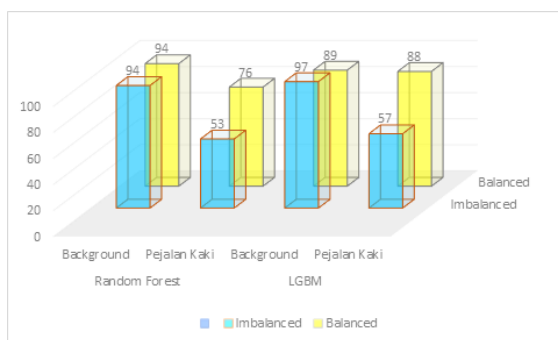
#### 4.3 Analisa Hasil

Berdasarkan eksperimen yang telah dilakukan dari 2 skenario, yaitu implementasi metode *ensemble learning* dan 2 blok layer conv dari arsitektur VGG-16 sebagai *feature extractor* pada dataset pedestrian dengan 1 objek pejalan kaki dihasilkan hasil segmentasi dari model yang telah dibuat dan akurasi setiap kelas serta nilai IoU pada setiap kelasnya. Pada bagian ini akan dijelaskan Analisa hasil dari perbandingan untuk skenario 1 dan 2 dari segi akurasi pada algoritma dan akurasi tiap kelas untuk tahap pengujian.

Pada skenario pertama menggunakan data yang belum diseimbangkan kelasnya (*imbalanced*) dan skenario kedua dengan data yang sudah diseimbangkan (*balanced*), grafik perbandingan kedua algoritma pada kedua skenario ditampilkan pada gambar 4



Gambar 4. Grafik perbandingan Imbalanced vs balanced dataset



Gambar 5. Grafik perbandingan tahap pengujian dengan Imbalanced vs balanced dataset

Pada gambar 4 ditampilkan grafik perbandingan dari kedua skenario yaitu pada kelas pejalan kaki dengan kedua algoritma mengalami peningkatan akurasi setelah proses *balanced data*. Grafik bar dengan warna biru adalah performa akurasi pada setiap kelas untuk *imbalanced data* pada kedua algoritma, sedangkan grafik bar dengan warna kuning untuk data yang sudah diseimbangkan. Pada algoritma random forest akurasi pejalan kaki mengalami peningkatan sebesar 31 %, sedangkan pada algoritma LGBM mengalami peningkatan sebesar 47 %. Pada tahap pengujian model dengan

data baru juga akan ditampilkan grafik perbandingan dari skenario kedua, yaitu dengan data yang seimbang yang akan ditunjukkan pada gambar 5.

Pada gambar 5 ditampilkan grafik perbandingan dari kedua skenario yaitu pada tahap pengujian model dengan data baru. Pada tahap ini dengan algoritma random forest akurasi pejalan kaki mengalami peningkatan sebesar 23 %, sedangkan dengan LGBM untuk kelas pejalan kaki mengalami peningkatan sebesar 31 %.

#### 5 KESIMPULAN

Berdasarkan pengujian yang telah dilakukan dalam perbandingan *imbalanced* dan *balanced data* pada segmentasi semantik gambar pedestrian single object menggunakan *ensemble learning* (EL) dan teknik oversampling sintetis dengan SMOTE dapat disimpulkan :

1. Menggunakan metode oversampling dengan teknik sintetis dapat meningkatkan akurasi pada metode EL untuk mengatasi masalah ketidakseimbangan kelas pada gambar pedestrian.
2. Pada proses pengujian dengan data baru, model yang dihasilkan dari dua algoritma EL tanpa class balancing menghasilkan gambar segmentasi pedestrian yang kurang akurat sebagai class minoritas dibandingkan dengan gambar groundtruthnya sebagai class mayoritas. Berdasarkan eksperimen yang telah dilakukan, perbandingan akurasi background dan object pedestrian terpaut jauh dan terdapat sedikit noise pada hasil segmentasi background
3. Pada pengujian model dengan class balancing, menghasilkan hasil segmentasi gambar pedestrian yang lebih akurat dibandingkan tanpa class balancing. Peningkatan akurasi pada object pedestrian cukup signifikan. Hal ini membuktikan oversampling dapat meningkatkan akurasi class minoritas pada segmentasi gambar, tapi dengan bertambahnya banyak data pada class minoritas menimbulkan noise (warna putih) yang berlebihan atau lebih banyak daripada tanpa class balancing pada background yang menyebabkan turunnya akurasi background.

#### DAFTAR PUSTAKA

- FERNÁNDEZ, A., GARCÍA, S., GALAR, M., PRATI, R. C., KRAWCZYK, B., & HERRERA, F., 2018. Learning from Imbalanced Data Sets. In *Learning from Imbalanced Data Sets*. [https://doi.org/10.1007/978-3-319-98074-4\\_5](https://doi.org/10.1007/978-3-319-98074-4_5)
- GUO, C., MA, Y., XU, Z., CAO, M., & YAO, Q., 2019. An Improved Oversampling Method for imbalanced Data-SMOTE Based on Canopy and K-means. *Proceedings - 2019 Chinese Automation Congress, CAC 2019*, 1467–1469.

- <https://doi.org/10.1109/CAC48633.2019.8997367>
- HASHEMI, S. R., SALEHI, S. S. M., ERDOGMUS, D., PRABHU, S. P., WARFIELD, S. K., & GHOLIPOUR, A., 2019. Asymmetric Loss Functions and Deep Densely-Connected Networks for Highly-Imbalanced Medical Image Segmentation: Application to Multiple Sclerosis Lesion Detection. *IEEE Access*, 7, 1721–1735.  
<https://doi.org/10.1109/ACCESS.2018.2886371>
- HE, H., & MA, Y., 2013. *Imbalanced Learning - Foundations, Algorithms and Applications*. 216.
- LIANG, M., CHANG, Z., WAN, Z., GAN, Y., SCHLANGEN, E., & ŠAVIJA, B., 2022. Interpretable Ensemble-Machine-Learning models for predicting creep behavior of concrete. *Cement and Concrete Composites*, 125(October 2021).  
<https://doi.org/10.1016/j.cemconcomp.2021.104295>
- LITJENS, G., KOOI, T., BEJNORDI, B. E., ARINDRA, A., SETIO, A., CIOMPI, F., GHAFORIAN, M., LAAK, J. A. W. M. VAN DER, GINNEKEN, B. VAN, & SÁNCHEZ, C. I., 2017. *A survey on deep learning in medical image analysis*. 42(December 2012), 60–88.  
<https://doi.org/10.1016/j.media.2017.07.005>
- LIU, Z., & WU, D., 2019. Unsupervised Ensemble Learning for Class Imbalance Problems. *Proceedings 2018 Chinese Automation Congress, CAC 2018*, 3593–3600.  
<https://doi.org/10.1109/CAC.2018.8623590>
- LU, M., & LI, F., 2020. Survey on lie group machine learning. *Big Data Mining and Analytics*, 3(4), 235–258.  
<https://doi.org/10.26599/BDMA.2020.9020011>
- NUGROHO, A., FANANI, A. Z., & SHIDIK, G. F., 2021. Evaluation of Feature Selection Using Wrapper for Numeric Dataset with Random Forest Algorithm. *Proceedings - 2021 International Seminar on Application for Technology of Information and Communication: IT Opportunities and Creativities for Digital Innovation and Communication within Global Pandemic, ISemantic 2021*, 179–183.  
<https://doi.org/10.1109/iSemantic52711.2021.9573249>
- RUANGTHONG, P., & JAIYEN, S., 2016. Hybrid ensembles of decision trees and Bayesian network for class imbalance problem. *2016 8th International Conference on Knowledge and Smart Technology, KST 2016*, 39–42.  
<https://doi.org/10.1109/KST.2016.7440523>
- SHUMALY, S., NEYSARYAN, P., & GUO, Y., 2020. Handling Class Imbalance in Customer Churn Prediction in Telecom Sector Using Sampling Techniques, Bagging and Boosting Trees. *2020 10th International Conference on Computer and Knowledge Engineering, ICCKE 2020*, 82–87.  
<https://doi.org/10.1109/ICCKE50421.2020.9303698>
- SMALL, H., & VENTURA, J., 2017. *Handling Unbalanced Data in Deep Image Segmentation*. <https://svds.com/learning-imbalanced-classes/>
- SYAHRIR, M., KUNCORO, A., ARI KURNIAWAN, A., ARIEF SOELEMEN, M., & FAJAR SHIDIK, G., 2018. Image Enhancement Segmentation and Edge Detection in MRI for Mammogram Disease. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, ISemantic 2018*, 515–521.  
<https://doi.org/10.1109/ISEMANTIC.2018.8549801>
- TAUBERT, O., GOTZ, M., SCHUG, A., & STREIT, A., 2020. Loss Scheduling for Class-Imbalanced Image Segmentation Problems. *Proceedings - 19th IEEE International Conference on Machine Learning and Applications, ICMLA 2020*, 426–431.  
<https://doi.org/10.1109/ICMLA51294.2020.00073>
- XIA, X., LU, Q., & GU, X., 2019. Exploring An Easy Way for Imbalanced Data Sets in Semantic Image Segmentation. *Journal of Physics: Conference Series*, 1213(2).  
<https://doi.org/10.1088/1742-6596/1213/2/022003>
- YANG, F., MA, Z., & XIE, M., 2021. Image Classification with Superpixels and Feature Fusion Method. *Journal of Electronic Science and Technology*, 19(1), 70–78.  
<https://doi.org/10.1016/j.jnlest.2021.100096>