UNIVERSITAS INDONESIA

PENGGUNAAN BOBOT DAN JARAK DALAM *FEATURE-TRANSFER LEARNING* UNTUK KLASIFIKASI GAMBAR

RINGKASAN DISERTASI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Doktor Ilmu Komputer

NINA SEVANI
1806261471

FAKULTAS ILMU KOMPUTER
PROGRAM DOKTOR ILMU KOMPUTER
DEPOK 2023

ABSTRAK

Nama : NINA SEVANI
Program Studi : Doktor Ilmu Komputer
Judul : PENGGUNAAN BOBOT DAN JARAK DALAM *FEATURE-TRANSFER LEARNING* UNTUK KLASIFIKASI GAMBAR
Pembimbing : Prof. Dr. Eng. Wisnu Jatmiko, S.T., M.Kom.

Transfer learning merupakan pengembangan dari pembelajaran mesin biasa (tradisional) yang dapat diterapkan pada *cross-domain*. *Cross-domain* adalah domain yang memiliki perbedaan pada *feature space* atau pada marginal dan *conditional distribution*, sehingga sulit ditangani dengan metode pembelajaran mesin biasa. Perbedaan ini banyak terjadi pada kasus *computer vision* atau *pattern recognition* seperti untuk mengenali korban bencana alam melalui foto yang diambil dari atas menggunakan *drone* atau helikopter. Terjadinya perbedaan *feature space* dan distribusi data ini karena adanya perbedaan sudut, cahaya, dan alat yang berbeda. Kondisi seperti ini semakin menyulitkan untuk dilakukannya klasifikasi gambar terlebih pada *domain* dengan keterbatasan label. Implementasi *transfer learning* terbukti dapat memberikan *performance* yang baik pada banyak kasus, termasuk kasus yang menggunakan dataset gambar.

Dalam *transfer learning* penting untuk menghindari terjadinya *negative transfer learning*, sehingga perlu dilakukan pengukuran kesamaan (*similarity*) antar domain. Penelitian ini menerapkan *feature-representation-transfer* dan menggunakan *Maximum Mean Discrepancy* (MMD) untuk mengukur jarak antar *feature* pada domain yang terlibat di *transfer learning*. Setelah mengukur kesamaan antar domain, maka akan dilakukan pemilihan *feature* berdasarkan jarak antar *feature*. *Feature* terpilih adalah *feature* yang mempunyai jarak kurang dari *threshold* yang ditentukan. Bobot akan diberikan kepada *feature* terpilih. Selain melakukan pemilihan *feature* berdasarkan kesamaan domain, metode ini juga melakukan pemilihan *feature* yang signifikan antar *class label* dan dalam *class label* dengan menggunakan ANOVA (*Analysis of Variance*). Hanya *feature* yang signifikan yang akan digunakan untuk proses prediksi.

Metode yang diusulkan juga menerapkan *inter-cluster class label* untuk memperkecil perbedaan *conditional distribution*. Prinsip kerja *inter-cluster class label* ini adalah menghitung jarak minimal dari *instance* pada domain target ke setiap *center of cluster class label*. Rumus jarak yang digunakan adalah *Euclidean Distance*. Properti statistik seperti rata-rata dan varians akan digunakan pada metode ini, untuk menggambarkan struktur data lokal dalam setiap domain. Penggunaan rata-rata digunakan untuk menentukan *threshold* dan pusat *cluster class label*, sedangkan varians digunakan untuk pemilihan *feature* yang signifikan. Proses prediksi label dilakukan berdasarkan *feature* terpilih yang telah diberi bobot dan jarak terpendek setiap *instance* ke salah satu *class* label.

Tidak terdapat parameter tambahan dalam fungsi pembelajaran yang diusulkan. Selain itu, proses penentuan label juga dilakukan tanpa iterasi, sehingga memungkinkan metode ini dapat dijalankan dengan keterbatasan *resource*. Hasil eksperimen menunjukkan bahwa metode yang diusulkan dapat memberikan *performance* sebesar 46,6%, pada saat menggunakan SVM sebagai *classifier* dan 51.7% pada saat menggunakan *logistic regression*. Akurasi yang didapat dengan SVM ini mengimbangi metode *feature-representation-transfer* sebelumnya. Namun akurasi dari *logistic regression* sudah dapat mengungguli metode sebelumnya. Hasil ini menunjukkan bahwa penggunaan metode *feature selection* menggunakan properti statistik yang dikombinasikan dengan pemberian bobot pada *feature* terpilih dan jarak minimal dapat memberikan hasil akurasi yang baik tanpa memerlukan *resource* yang besar.

Kata Kunci:
feature-representation-transfer, MMD, bobot, rata-rata, jarak.

ABSTRACT

Name : NINA SEVANI
Study Program : Doctor of Computer Science
Title : THE UTILISATION OF WEIGHT AND DISTANCE IN *FEATURE*-TRANSFER LEARNING FOR IMAGE CLASSIFICATION
Counsellor : Prof. Dr. Eng. Wisnu Jatmiko, S.T., M.Kom.

Transfer learning is the extension of traditional machine learning in a cross-domain environment. Cross-domains are domains with different feature spaces or different marginal and conditional distributions. Many real-world cases of computer vision and pattern recognition, such as the surveillance of some victims of natural disasters from above using a drone or helicopter, have these differences. These conditons are difficult to handle with traditional machine learning methods. The differences in feature space or data distribution caused by the existence of different angles, different light, and different tools. All of these situation add difficulty to the classification process, especially in domains with limited labels. The implementation of transfer learning is proven to provide good performance in many cases of cross-domain learning, including cases that use image datasets.

In transfer learning, it is important to measure the similarity between domains to avoid negative transfer learning. This study applies feature-representation-transfer and uses Maximum Mean Discrepancy (MMD) to measure the distance between features in the cross-domains and reduce the domain discrepancy. After measuring the similarity between domains, a feature selection will be made based on the distance between the features. Selected features are features that have a distance less than the specified threshold. Weight will be given to the selected features. In addition to selecting features based on domain similarity, this method also selects significant features between class labels and within class labels using ANOVA (Analysis of Variance). Only significant features will be used for the prediction process.

The proposed method also applies an inter-cluster class label to minimize the difference in conditional distribution. The inter-cluster class label works by calculating the minimum distance from the instance in the target domain to each center of the cluster class label. The distance formula used is Euclidean distance. Statistical properties such as mean and variance will be used in this method to describe the local data structure in each domain. The average is used to determine the threshold and center of the cluster class label, while the

variance is used to select significant features. The label prediction process is carried out based on the selected features that have been weighted and the shortest distance for each instance to one of the label classes.
There are no additional parameters in the proposed learning function. In addition, the process of determining the label is also carried out without iteration, thus allowing this method to be run with limited resources. The experimental results show that the proposed method can provide a performance of 46.6% when using SVM as a classifier and 51.7% when using logistic regression. The accuracy obtained from SVM offsets the previous feature-representation transfer learning. However, the accuracy of logistic regression has been able to outperform the previous method. These results indicate that the use of the feature selection method using statistical properties combined with assigning weights to selected features and a minimum distance can provide good accuracy without requiring large resources.

Key words:
feature-representation-transfer, MMD, weight, *mean*, distance

DAFTAR ISI

ABSTRAK i
ABSTRACT iii
DAFTAR ISI v
DAFTAR GAMBAR vii
DAFTAR TABEL viii
DAFTAR ISTILAH ix
DAFTAR NOTASI xi
BAB 1 PENDAHULUAN 1
1.1. Latar Belakang 1
1.2. Rumusan Masalah 5
1.3. Tujuan 5
1.4. Metodologi Penelitian 6
1.5. Batasan Masalah 8
BAB 2 TRANSFER LEARNING 10
2.1 Perbedaan Pembelajaran Mesin dengan Transfer Learning 10
2.2 Cross-domain 12
2.3 Pengelompokkan Transfer Learning 12
2.4 Komponen dalam Transfer Learning 16
2.5 Beragam Teknik Pembelajaran pada Transfer Learning 17
2.6 Feature-Representation Transfer (Feature-based Transfer Learning) 19
BAB 3 METODOLOGI PENELITIAN 23
3.1. Gambaran Umum WbFTL 23
3.2. *Feature* Transformasi 1 dengan Pemilihan *Feature* (*Feature* Selection) menggunakan *Threshold* 26
3.3. Transformasi *Feature* 2 dengan Teknik Pemilihan *Feature* Menggunakan ANOVA 28
3.4. Transformasi *Feature* 3 dengan *Inter-Cluster Class* Label 29
3.5. Resume Cara Kerja WbFTL 30

BAB 4 HASIL PENELITIAN 32
4.1. Dataset dan Eksperimen Setting 32
4.2. Metode State-of The-Art 33
4.3. Algoritma WbFTL 33
4.4. Hasil Eksperimen dengan Classifier SVM 35
4.5. Hasil Eksperimen dengan Classifier Logistic Regression 37
4.6. Perbandingan Hyperparameter 38
BAB 5 PENUTUP 39
5.1. Kesimpulan 39
5.2. Kelebihan Metode WbFTL 40
5.3. Kekurangan Metode WbFTL 41
5.4. Keterbatasan Metode WbFTL 41
5.5. Saran Pengembangan 41
Daftar Pustaka 43

DAFTAR GAMBAR

Gambar 1.1. Tahapan Metode 8

Gambar 2.1. Ilustrasi Perbedaan Pembelajaran Mesin Biasa dengan Transfer Learning 11

Gambar 2.2. Pengelompokkan Transfer Learning 14

Gambar 2.3. Tahapan Umum Pendekatan dengan Iterasi pada Feature-based Transfer Learning 21

Gambar 2.4. Tahapan Umum Pendekatan tanpa Iterasi pada Feature-based Transfer Learning 22

Gambar 3.1. Gambaran Umum WbFTL 24

Gambar 3.2. Proses Training dan Inference pada WbFTL 27

Gambar 3.3. Tahapan Transformasi Feature pada WbFTL 28

Gambar 3.4. Ilustrasi Perbedaan Menyeluruh antara WbFTL dengan Metode Feature-based Transfer Learning Lainnya 31

Gambar 4.1. Contoh Gambar Asli Dataset Amazon 33

Gambar 4.2. Perbandingan Akurasi WbFTL dengan Easy Feature-based Transfer Learning per Pasangan Dataset 335

Gambar 4.3. Perbandingan Akurasi WbFTL dengan Iterative Feature-based Transfer Learning per Pasangan Dataset 36

DAFTAR TABEL

Tabel 2.1. Pendekatan pada Transfer Learning..... 15

Tabel 2.2. Penelitian-penelitian tentang Pendekatan Transfer Learning 15

Tabel 2.3. Penelitian tentang Beragam Teknik pada Transfer Learning..... 17

Tabel 3.1. Resume Perbandingan WbFTL dengan Metode Feature-based Transfer Learning Lainnya 30

Tabel 4.1. Penjelasan Dataset untuk WbFTL 32

Tabel 4.2. Algoritma WbFTL 33

Tabel 4.3. Perbandingan Akurasi WbFTL dengan Easy Feature-based Transfer Learning 35

Tabel 4.4. Perbandingan Akurasi WbFTL dengan Iterative Feature-based Transfer Learning 36

Tabel 4.5. Perbandingan Akurasi WbFTL dengan Logistic Regression pada Easy Feature-based Transfer Learning 37

Tabel 4.6. Perbandingan Akurasi WbFTL dengan Logistic Regression pada Iterative Feature-based Transfer Learning 38

Tabel 4.7. Perbandingan Jumlah Parameter WbFTL dengan Metode Feature-based Transfer Learning Sebelumnya 38

DAFTAR ISTILAH

Singkatan		Kepanjangan	Arti
ANOVA	:	*Analysis of Variance*	Metode statistik untuk menguji perbedaan rata-rata antar kelompok.
MMD	:	*Maximum Mean Discrepancy*	Pengukuran probabilitas (distribusi data) yang menggunakan persamaan jarak pada suatu ruang.
RKHS	:	*Reproducing Kernel Hilbert Space*	Sebuah fungsi kernel yang akan memetakan titik asli ke titik pada suatu ruang Hilbert, yang mana proses evaluasi titik menggunakan fungsi linear kontinyu.
SSB	:	*Sum of Squares Between*	Jumlah kuadrat antar kelompok, yang mana bila semua kelompok sama rata-ratanya maka SSB akan bernilai 0.
SSW	:	*Sum of Squares Within*	Jumlah kuadrat dalam kelompok.
SURF	:	*Speeded-Up Robust Features*	Salah satu metode *local feature detector* dan *feature descriptor*.
SVM	:	*Support Vector Machines*	Salah satu *classifier* yang bersifat non-parametric dengan cara kerja membentuk *soft margin*.
WbFTL	:	*Weighted-based Feature Transfer Learning*	Nama metode *feature-transfer learning* yang menekankan proses transformasi *feature* secara mudah dengan menerapkan pendekatan pemilihan dan pembobotan *feature* digabung dengan properti statistik dan jarak.
LR	:	*Logistic Regression*	Salah satu *classifier* yang bersifat non-parametric dengan cara kerja menghitung nilai *probability*.
1-NN	:	*1-Nearest Neighbor*	Salah satu *classifier* yang bersifat parametric dengan cara kerja mencari *class*

			terdekat yang biasanya dilakukan berdasarkan persamaan jarak.
ViT	:	*Vision Transformer*	Salah satu arsitektur *deep learning* yang menggabungkan arsitektur *convolutional* dengan arsitektur *vision* yang menggunakan mekanisme *attention* untuk mendapatkan *semantic feature* secara lebih baik (mudah dilatih, tidak membutuhkan sumber daya besar).
ED	:	*Euclidean Distance*	Perhitungan jarak dua buah titik pada ruang euclidean.

DAFTAR NOTASI

Notasi	Arti
D_S	Domain Asal
D_T	Domain Tujuan (Target)
n_S	Jumlah *instance* pada domain asal
n_T	Jumlah *instance* pada domain tujuan
$\mathcal{X}_S$	*Feature Space* pada domain asal
$\mathcal{X}_T$	*Feature Space* pada domain tujuan
x_{Si}	*Instance* ke-i dari domain asal
x_{Tj}	*Instance* ke-j dari domain target
nf_S	Jumlah *feature* pada domain asal
nf_T	Jumlah *feature* pada domain tujuan
b	Jumlah *feature* hasil transformasi menggunakan MMD dan kernel, jumlahnya sama dengan *feature* awal data sebelum menerapkan MMD.
f_{Si}	*Feature* ke-i dari domain asal
f_{Tj}	*Feature* ke-j dari domain tujuan
$\bar{f}_{Si}$	Rata-rata *Feature* ke$-i$, dari $\mathcal{X}_S$
$\bar{f}_{Ti}$	Rata-rata *Feature* ke$-i$, dari $\mathcal{X}_T$
F_m	Hasil transformasi *feature* untuk mengurangi perbedaan domain, menggunakan MMD dan kernel
f_{m_i}	*Feature* ke-*i* hasil transformasi menggunakan MMD dan kernel
Th	*Threshold* (batas yang digunakan untuk pemilihan *feature*)
F_z	Subset *feature* hasil penerapan nilai *threshold* pada f'
f_{zi}	*Feature* ke-*i* dari F_z
F_{Trans1}	Transformasi *feature* pertama
$F_{STrans1}$	Transformasi *feature* pertama untuk D_S, yaitu subset *feature* hasil penerapan nilai *threshold* pada f' yang telah diberi bobot.
$f_{STrans1(i)}$	Transformasi *feature* pertama untuk *instance* ke-i pada $F_{STrans1}$.
$F_{TTrans1}$	Transformasi *feature* pertama untuk D_T, yaitu subset *feature* hasil penerapan nilai *threshold* pada f' yang telah diberi bobot
$f_{TTrans1(i)}$	Transformasi *feature* pertama untuk *instance* ke-i pada $F_{TTrans1}$.
k	Jumlah *feature* pada subset *feature* F_{Trans1}
$w(f_{s_i})$	Bobot *feature* terpilih ke-i dari subset *feature* F_{Trans1}

a	Jumlah *feature* pada subset *feature* $F_{STrans2}$
F_{Trans2}	Transformasi *feature* kedua
$F_{STrans2}$	Transformasi *feature* kedua untuk D_S, yaitu subset *feature* hasil penerapan ANOVA pada $F_{STrans1}$
$f_{STrans2(i)}$	Transformasi *feature* untuk *instance* ke-i pada $F_{STrans2.}$
C	Jumlah class label
F_{Trans3}	Transformasi *feature* ketiga
$F_{TTrans3}$	Transformasi *feature* ketiga untuk D_T, yaitu subset *feature* hasil penerapan inter-cluster class label pada $F_{TTrans1}$
$f_{TTrans3(i)}$	Transformasi *feature* untuk *instance* ke-i pada $F_{TTrans3.}$
$d_{(i,j)}$	Jarak *instance* ke-i dari domain tujuan ke pusat class label j
$h_{c(yi)}$	Pusat *Cluster Class* Label c
y_T	*Predicted label* untuk *instances* pada domain tujuan
$dist(X_{Ti}, h_{C(yi)})$	Jarak minimal *instance* ke-i dari domain tujuan menuju pusat *cluster class* label j

BAB 1
PENDAHULUAN

1.1. Latar Belakang

Keberhasilan pembelajaran mesin dalam menyelesaikan beragam masalah kompleks hanya terbatas untuk kondisi yang mana data *training* dan data *testing* mempunyai distribusi yang sama atau berada pada domain yang sama dengan *feature space* yang sama pula. Pada kenyataannya banyak kasus nyata yang tidak berada pada domain sama atau tidak mempunyai kesamaan distribusi data. Sebagai contoh, dalam kasus prediksi korban bencana alam yang dilakukan berdasarkan foto yang diambil kamera jarak jauh yang dioperasikan oleh manusia melalui helikopter atau sebuah *drone* (Kurniawan, 2019). Pengambilan foto dengan jenis kamera yang berbeda, sudut pengambilan yang berbeda, kondisi lingkungan yang berbeda (misalkan pada suasana cerah, hujan, atau berkabut), posisi objek, bahkan prosedur pengambilan foto yang berbeda bisa menghasilkan distribusi data yang berbeda juga (Cheplygina et al., 2018). Pada kasus seperti ini, penggunaan algoritma pembelajaran mesin biasa akan sulit diterapkan, disamping kemungkinan keberhasilan yang juga kecil.

Mengingat keterbatasan dalam algoritma pembelajaran mesin biasa ini, maka diperlukan *knowledge transfer* atau *transfer learning* (Pan & Yang, 2010; Weiss et al., 2016a; Shao et al., 2015; Day & Khoshgoftaar, 2017; Zhuang et al., 2019). *Transfer learning* adalah salah satu metodologi pembelajaran mesin dengan fokus melakukan *transfer knowledge* antar domain yang berbeda namun berkaitan (*cross-domain*) (Pan & Yang, 2010; Shao et al., 2015; Zhuang et al., 2019). Perbedaan domain disini ditekankan pada perbedaan *feature space* atau perbedaan distribusi data yang bisa terjadi karena perbedaan alat, prosedur, dan bagian pada saat pengumpulan data (Wang, Chen, et al., 2018; Cheplygina et al., 2018). Dengan adanya kesamaan atau perbedaan *feature space* antar domain dan ketersediaan *class* label pada domain asal maupun domain target, terdapat beberapa pendekatan yang dapat digunakan pada *transfer learning*. Pendekatan-pendekatan tersebut adalah *instance-transfer*, *feature-representation transfer*, *model-transfer*, dan *relational-knowledge-transfer* (Pan & Yang, 2010). Dalam implementasinya, setiap pendekatan tersebut menggunakan kombinasi dari beberapa teknik. Tidak ada panduan pasti untuk menentukan teknik terbaik untuk dapat menghasilkan akurasi terbaik. Penelitian terdahulu menampilkan bahwa akurasi dari metode

pembelajaran mesin biasa, seperti *Nearest Neighbor* dan *Principal Component Analysis* dalam kasus *cross-domain* masih sangat rendah hanya mencapai 31.4% dan 43.6% (Wang, Feng, et al., 2018b). Penerapan metode *transfer-learning* juga masih menunjukkan hasil yang rendah, walaupun lebih baik daripada metode pembelajaran mesin biasa.

Selain kemampuan untuk mengatasi kondisi *cross-domain*, *transfer learning* juga memungkinkan proses *learning* dilakukan dengan lebih efisien karena *knowledg*e dari domain asal akan ditransfer ke domain target, sehingga domain target tidak perlu lagi melakukan proses *training* dari awal (*learning from scratch*) dan dapat menghemat sumber daya yang ada. Domain asal adalah domain yang berisi data *training*, sementara domain target atau domain target adalah domain yang berisi data *testing* (Shao et al., 2015). Selain kemampuan untuk dapat diterapkan pada domain yang berbeda, metode *transfer learning* terutama yang termasuk dalam kelompok "mudah", juga terbukti dapat mengurangi kompleksitas *training* pada domain target sehingga memungkinkan untuk diimplementasikan pada kondisi dengan keterbatasan jumlah sumber daya (Wang et al., 2019). Selain itu, karena *knowledge* yang digunakan untuk proses pembelajaran (*training*) juga berasal dari data yang mengandung kemiripan dengan data yang akan diklasifikasikan (*testing*), maka diharapkan akurasi yang dihasilkan akan baik.

Karena kemampuannya dalam mentransfer *knowledge* antar domain, konsep *transfer learning* ini telah diimplementasikan dalam berbagai bidang. Bahkan beragam strategi, pendekatan, teknik, juga telah diusung oleh peneliti *transfer learning*. Meskipun penelitian tentang *transfer learning* terus berkembang dengan hasil yang makin baik, melalui penerapan beberapa teknik, namun pada dasarnya tidak ada panduan pasti akan kombinasi terbaik untuk menghasilkan akurasi yang tinggi. Sehingga penggunaan kombinasi dari beberapa teknik pada penelitian *transfer learning* untuk dapat menghasilkan akurasi yang baik masih merupakan *open problem* (Weiss et al., 2016a). Dari beragam kemungkinan kombinasi teknik yang dilakukan pada penelitian *transfer learning*, dapat dilihat bahwa pemanfaatan *instances* tidak berlabel untuk mengurangi perbedaan distribusi antar domain belum dimanfaatkan secara optimal. Padahal dengan mengingat bahwa salah satu kunci keberhasilan dalam adaptasi domain yang berbeda adalah pada *instance* khususnya di domain target (*instance-based adaptation*) dan pada *feature matching (feature-based adaptation)* (Wang, Feng, et al., 2018b). Pada *feature-based adaptation* diasumsikan bahwa terdapat *common space* yang

mana pada *cross-domain*. Asumsi ini sendiri lebih ringan daripada asumsi pada *instance-based adaptation* (J. Zhang et al., 2017). Selain itu, penelitian tentang metode yang dapat beradaptasi dengan dataset dan melakukan *feature discovery* juga masih terbuka (Li et al., 2020). Oleh karena itu penelitian ini akan berfokus pada *feature-representation transfer learning* yang akan melakukan transformasi *feature* untuk mengatasi perbedaan distribusi pada *cross-domain*.

Selain itu, diantara beragam pendekatan pada *transfer learning*, *feature-representation transfer* atau *feature-based transfer learning* merupakan pendekatan yang banyak digunakan karena dapat diimplementasikan pada beragam bidang (Pan & Yang, 2010; Zhuang et al., 2021; Day & Khoshgoftaar, 2017) dengan beragam kombinasi teknik dan teknologi (Gong et al., 2012; Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Sanodiya et al., 2020; Long, Wang, Ding, Sun, et al., 2014; Long et al., 2013a; Tahmoresnezhad & Hashemi, 2017a; Wang, Feng, et al., 2018b; Wang et al., 2020; Wang et al., 2019; Sun et al., 2016; Fernando et al., 2013). Salah satu bentuk implementasi yang cukup banyak adalah untuk klasifikasi gambar (Gong et al., 2012; Long et al., 2013b; Wang, Feng, et al., 2018b; Wang et al., 2020; Wang et al., 2017; Long, Wang, Ding, Sun, et al., 2014; Sanodiya et al., 2020; Wang, Feng, et al., 2018b; Wang et al., 2019; Sun et al., 2016; Fernando et al., 2014; Tahmoresnezhad & Hashemi, 2017a). Penggunaan *feature-based transfer learning* untuk klasifikasi gambar meliputi beragam bidang, mulai dari medis khususnya yang terkait dengan gambar medis, *bioinformatics*, robotika, *computer vision*, dan *natural language processing* (Kouw & Loog, 2021).

Pada dasarnya terdapat beberapa strategi untuk melakukan transformasi *feature*, seperti *feature augmentation*, *feature alignment*, *feature clustering*, *feature encoding*, *feature mapping*, dan *feature selection*. Umumnya setiap strategi ini digunakan untuk kasus yang berbeda (Zhuang et al., 2021). Seperti contohnya *feature clustering* digunakan pada kasus pemrosesan teks atau dokumen, *feature encoding* digunakan untuk kasus yang menggunakan *deep learning*. Penelitian terdahulu tentang *feature-representation transfer* lebih banyak menggunakan strategi *feature mapping* untuk domain gambar. Penerapan *feature mapping* ini dilakukan dengan menggunakan pendekatan pengurangan dimensi dan membentuk matrik proyeksi. Penerapan strategi pemilihan *feature* (*feature selection*) pada penelitian *feature-representation transfer* masih jarang, sehingga masih terbuka untuk ditelaah lebih dalam.

Pemilihan *feature* sendiri juga melakukan pengurangan dimensi dengan cara yang lebih sederhana. Pemilihan *feature* juga memungkinkan untuk menghasilkan sistem dengan unjuk kerja yang baik, karena fokus pada *feature* yang mempunyai tingkat relevansi tinggi untuk menghasilkan label pada kasus klasifikasi (Kouw & Loog, 2021; Deng et al., 2019). Selain itu pemilihan *feature* juga tidak tergantung pada algoritma pembelajaran yang dihasilkan, sehingga lebih fleksibel dalam penerapannya (Pham et al., 2020). Berdasarkan nilai akurasi yang didapat oleh metode terbaik pada setiap kelompok tersebut maka masih memungkinkan untuk ditingkatkan dengan kombinasi teknik yang tepat, sehingga dapat dijalankan secara efisien dengan memberikan hasil akurasi yang baik. Selain itu kedua pendekatan tersebut juga kurang memperhatikan hubungan antara *feature* dengan informasi label, baik pada D_S maupun D_T. Kedua kondisi inilah yang memotivasi untuk dilakukannya penelitian pada *feature-representation transfer learning*.

Classifier umum digunakan dalam kasus *feature-representation transfer learning*. Penggunaan *classifier* ini akan dikombinasikan dengan metode transformasi *feature* yang diusung oleh masing-masing peneliti. *Classifier* akan berfungsi untuk melakukan klasifikasi, dengan *input* data yang sudah mengalami transformasi *feature*.

Dalam rangka melakukan *benchmarking* dengan *state-of-the-art* metode *feature-representation transfer learning*, maka *classifier* yang digunakan juga akan dibuat sama, agar perbandingan dilakukan secara adil. *Classifier* 1-NN dan SVM akan digunakan pada penelitian awal. Penelitian awal ditujukan untuk melihat unjuk kerja kedua *classifier*. *Classifier* dengan hasil yang lebih baik untuk dataset gambar akan digunakan sebagai *classifier* utama dalam penelitian yang dilakukan.

Selain kedua *classifier* tersebut, penelitian awal tentang *transfer learning* juga menerapkan *classifier logistic regression* (Y. Zhang et al., 2010). *Classifier logistic regression* dipilih karena prinsip kerjanya yang sederhana dan dapat digunakan untuk kasus klasifikasi dengan banyak label (*multi-class classification*) menggunakan pendekatan *transfer learning*. Hal ini juga yang memotivasi digunakannya *logistic regression* sebagai salah satu classifier untuk penelitian yang dilakukan.

Berdasarkan pada apa yang sudah dilakukan pada penelitian tentang *feature-representation transfer learning* sebelumnya, maka penelitian ini mengusulkan konsep *feature-transfer*

learning yang disebut sebagai *Weighted-based Feature Transfer Learning* (WbFTL). WbFTL mengusung konsep "sederhana" yang dapat dilihat dari tidak digunakannya parameter dalam proses transformasi *feature*. Sehingga tidak perlu menggunakan sumber daya yang besar dalam proses implementasinya. Terdapat tiga tahapan transformasi *feature* pada WbFTL, dua tahapan pertama menggunakan pendekatan pemilihan *feature* dan satu tahapan menggunakan pendekatan jarak. Pemilihan *feature* dilakukan menggunakan metode filter yang menggunakan properti statistik yaitu rata-rata sebagai *threshold* untuk menghasilkan representasi *feature* pertama. Dilanjutkan dengan pemilihan *feature* menggunakan ANOVA (*Analysis of Variance*) untuk menghasilkan representasi *feature* kedua. Pada transformasi *feature* kedua ini akan dipilih *feature* yang signifikan terhadap label, karena pada transformasi *feature* pertama tidak memperhatikan informasi label. Transformasi kedua akan digunakan sebagai masukan bagi *classifier* yang digunakan. *Classifier* ini juga akan digunakan untuk melakukan inferensi pada transformasi *feature* ketiga. Transformasi *feature* ketiga dilakukan menggunakan pendekatan jarak. Hasil proses inferensi menggunakan transformasi *feature* ketiga adalah label hasil prediksi.

1.2. Rumusan Masalah

Berdasarkan isu penting dan juga permasalahan yang diketahui berdasarkan latar belakang dan penelitian sebelumnya, maka dirumuskan tiga pertanyaan penelitian yang akan dibahas, yaitu:

1. "Apakah pendekatan pemilihan *feature* dan pembobotan *feature* dalam *feature-based transfer learning* dapat memberikan hasil klasifikasi yang baik?"
2. "Apakah penerapan *feature mapping* untuk *cross-domain* gambar dapat menghasilkan akurasi yang baik untuk kasus klasifikasi?"
3. "Apakah penggunaan *feature-based transfer learning* dengan minimal parameter dapat menghasilkan akurasi klasifikasi yang baik?"

1.3. Tujuan

Tujuan penelitian ini adalah menerapkan kombinasi dari beberapa teknik seperti MMD, bobot, jarak, dan properti statistik yaitu rata-rata dan varian untuk menghasilkan metode *feature-representation transfer* yang efisien dan hasil akurasi yang baik.

Inti dari metode *feature-representation transfer* adalah melakukan transformasi *feature* untuk membentuk representasi *feature* baru yang akan digunakan oleh model atau *classifier* untuk menghasilkan label hasil prediksi. Transformasi *feature* dilakukan dengan kombinasi beberapa teknik, yaitu *feature mapping* menggunakan MMD dan kernel, *feature selection* dan pemberian bobot pada *feature* terpilih menggunakan properti statistik (rata-rata dan varian), dan jarak untuk menghasilkan label hasil prediksi (label bagi domain tujuan). Diharapkan dengan menerapkan beberapa teknik dalam proses transformasi *feature*, maka ketiga isu pada *penelitian feature-representation transfer* dapat diatasi. Sehingga dapat dihasilkan metode *feature-representation transfer* yang efisien (tanpa banyak parameter) dengan hasil akurasi yang baik.

1.4. Metodologi Penelitian

Berdasarkan penjelasan tentang *feature-based transfer learning* diatas, maka penelitian ini mengusulkan penggunaan kombinasi beberapa cara untuk menyelesaikan tiga isu, yaitu mengurangi distribusi data dan upaya untuk meningkatkan unjuk kerja sistem tanpa menggunakan banyak parameter. MMD yang digabung dengan kernel serta pemanfaatan *instances* tidak berlabel pada domain target, merupakan cara untuk mengurangi perbedaan domain antar domain. Sementara peningkatan unjuk kerja dilakukan dengan menerapkan teknik seleksi *feature* dan pemberian bobot pada *feature* terpilih. Dalam metode *feature-representation transfer learning* ini juga akan menggunakan properti statistik, yaitu rata-rata dan varian dari data yang diolah, disamping juga menggunakan jarak yang akan diterapkan dalam bentuk *Euclidian Distance*.

Penelitian terdahulu mengungkapkan beberapa metode dan teknik untuk mengatasi isu-isu tersebut, mulai dari teknik sederhana untuk mencari kesamaan antar domain (Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Pan et al., 2008a), penambahan bobot selain menghitung kesamaan antar domain (Zhong et al., 2018), sampai menggunakan struktur data lokal dalam proses transfer pengetahuan (Wang, Chen, et al., 2018; Wang et al., 2019). Penambahan bobot selain dapat mengurangi perbedaan distribusi data antar domain (Dai & Xue, n.d.; Xu et al., 2017) juga dapat mengatasi masalah ketidakseimbangan jumlah *instance* antar domain (Wang et al., 2017), sehingga pada akhirnya juga akan dapat meningkatkan unjuk kerja dari *transfer learning* itu sendiri (Xia et al., 2015; Zhong et al., 2018). Penggunaan struktur data lokal

dalam proses transfer pengetahuan dapat mengurangi sumber daya (Wang et al., 2019), selain juga dapat mengurangi perbedaan distribusi data antar domain. Salah satu struktur data lokal yang banyak digunakan adalah rata-rata dan varian sebagai properti statistik, yang terbukti dapat memberikan unjuk kerja baik dalam proses klasifikasi (Wang et al., 2019; Sevani et al., 2019). Selain itu rata-rata juga merupakan pengukuran statistik yang menyatakan *central tendency* sehingga dapat menggambarkan seluruh distribusi data menjadi satu nilai tunggal (Manikandan, 2011; Frederick J Gravetter and Larry B. Wallnau, 2013).

Khusus untuk mengatasi permasalahan khusus pada *feature-representation transfer*, yaitu tentang bagaimana melakukan transformasi *feature*, penelitian ini juga menggunakan pendekatan *feature mapping* sebagai strategi pembentukan representasi *feature* yang baru. Fungsi pembelajaran yang digunakan pada penelitian ini juga akan menggabungkan antara properti statistik, jarak, dan juga bobot. Rata-rata (*mean*) dan varian akan menjadi properti statistik yang digunakan dalam setiap tahapan pembelajaran yang dilakukan, mulai untuk mengurangi perbedaan distribusi sampai dengan untuk proses transformasi *feature*. Mengingat distribusi yang sebenarnya sulit diketahui sehingga perlu pendekatan perkiraan yang dapat menjelaskan tentang data secara keseluruhan dan adaptif terhadap perubahan jumlah ataupun nilai data (Frederick J Gravetter and Larry B. Wallnau, 2013; Maity, 1977).

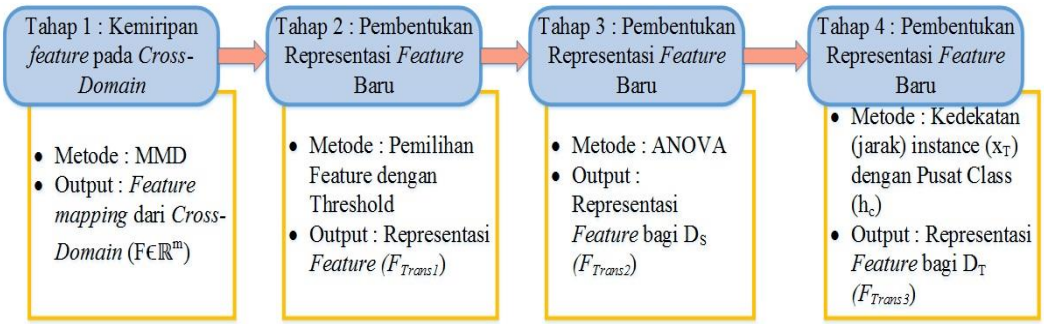
Penerapan *feature mapping* dilakukan untuk mengurangi perbedaan *marginal distribution* pada *cross-domain*. Pada tahapan ini, pembelajaran dilakukan dengan membentuk representasi *feature* yang baru menggunakan konsep kernel pada rumus MMD (Borgwardt et al., 2006). Fungsi mapping yang digunakan adalah fungsi jarak yang diukur menggunakan *Euclidean Distance* dari setiap pasangan *feature* rata-rata (*features mean*) pada *cross-domain*.

Sementara untuk mengurangi perbedaan *conditional distribution*, digunakan cara *inter-cluster class label distance*. Cara ini adalah mengukur jarak *instance* pada domain target ke masing-masing pusat *cluster class label*. Dari hasil perhitungan ini kemudian diambil jarak terpendek yang akan digunakan dalam penentuan label bagi *instance* yang bersangkutan.

Tahapan pembelajaran yang digunakan untuk meningkatkan unjuk kerja sistem (akurasi dan minimal parameter) adalah penerapan seleksi *feature* dan pemberian bobot pada *feature* yang terpilih. Pemberian bobot dianggap dapat meningkatkan unjuk kerja sistem dalam penentuan label (Zhong et al., 2018; Van Opbroek et al., 2019). Seleksi *feature* diterapkan dengan

menggunakan batas (*threshold*) berdasarkan properti statistik pada representasi *feature* yang baru, yang terbentuk dari hasil pembelajaran pengurangan perbedaan distribusi yang dilakukan pada tahap sebelumnya. Selain menggunakan *threshold* untuk menyeleksi *feature* yang akan digunakan, *feature* juga akan dipilih menggunakan ANOVA. ANOVA akan memilih *feature* yang signifikan dalam setiap *class* label dan juga antar *class* label. Penggunaan ANOVA bertujuan untuk memilih *feature* yang relevan dengan memperhatikan informasi *class* label. Strategi pemilihan *feature* menggunakan ANOVA juga dilakukan dalam rangka untuk mengatasi perbedaan *conditional distribution* pada *cross-domain* dengan memanfaatkan informasi *class* label.

Gambar 1.1 berikut menjelaskan tentang tahapan penelitian yang dilakukan, terutama pada bagian pembelajaran (*learning*). Melalui Gambar 1.1 dapat dilihat metode untuk setiap tahap serta hasil proses yang dilakukan.



Gambar 1.1. Tahapan Metode

1.5. Batasan Masalah

Fokus pada penelitian ini adalah melakukan transformasi *feature* dengan menerapkan beberapa strategi, seperti pendekatan *feature mapping* dan pemilihan *feature* pada *cross-domain*. Pendekatan yang digunakan untuk tahap pembelajaran dirancang untuk dapat mengurangi perbedaan *marginal* dan *conditional distribution* pada *cross-domain* sekaligus melakukan peningkatan unjuk kerja dengan meminimalisir penambahan parameter pada proses pembelajaran. Sesuai dengan yang disampaikan pada rumusan permasalahan diatas, maka berikut merupakan batasan permasalahan yang akan dikaji dalam penelitian ini:

A. Dataset:

Dataset yang digunakan adalah dataset gambar dengan kategori *real-world object*. Untuk label yang digunakan akan dibatasi pada 10 label yang umum, seperti yang juga digunakan pada penelitian sebelumnya. Seluruh dataset gambar yang digunakan sudah diproses dengan algoritma SURF (*Speeded Up Robust Features*) sebagai algoritma *feature detector* dan *descriptor*.

B. *Classifier*:

Metode *feature-based transfer learning* yang digunakan pada penelitian ini menggunakan *classifier* SVM. Penelitian sebelumnya pada umumnya menggunakan *classifier* SVM dan 1-NN. Hasil eksperimen awal menunjukkan bahwa SVM memberikan hasil yang lebih baik pada saat digunakan untuk klasifikasi gambar pada *cross-domain*.

Sementara untuk pengembangan penelitian akan dilakukan menggunakan *classifier Logistic Regression*. Meskipun *Logistic Regression* pada dasarnya untuk klasifikasi biner (*binary classification*), namun dapat digunakan juga untuk *multi-class classification*. Strategi yang digunakan adalah *one-vs-rest* (OvR), karena cocok untuk dataset yang jumlah *class* label tidak terlalu banyak. Kondisi ini sesuai dengan kondisi dataset yang digunakan pada penelitian WbFTL ini.

C. Pengujian unjuk kerja:

Untuk menguji unjuk kerja metode *feature-based transfer learning* pada penelitian ini, maka digunakan pengukuran menggunakan *confusion* matrix, khususnya dengan mengukur akurasi. Hasil pengukuran akurasi metode *feature-based transfer learning* yang dihasilkan pada penelitian ini akan dibandingkan dengan *state-of-the-art* atau metode *feature-based transfer learning* lainnya yang juga menggunakan dataset sejenis.

BAB 2
TRANSFER LEARNING

Transfer learning lahir dari konsep pembelajaran mesin yang sudah ada sebelumnya. *Transfer learning* memungkinkan pembelajaran dilakukan antar domain yang memiliki perbedaan *feature set* atau distribusi data. Dengan menerapkan *transfer learning*, pengetahuan dari satu domain dapat ditransfer ke domain lain. Domain yang menjadi awal pembelajaran disebut sebagai domain asal. Pengetahuan yang diperoleh dari domain asal ini yang kemudian ditransfer dan digunakan sebagai dasar pembelajaran bagi domain target (domain tujuan). Dengan adanya pengetahuan yang ditransfer ini, maka proses pembelajaran pada domain target bisa menjadi lebih cepat karena tidak perlu melakukan pembelajaran dari awal. Dengan teknik dan metode tertentu, *transfer learning* memungkinkan proses transfer ini terjadi antar domain yang berbeda, yang memiliki perbedaan distribusi data atau perbedaan *feature space*. Pengggunaan *transfer learning* sendiri sudah dimulai sejak tahun 2000. Beragam metode dan pendekatan sudah dicoba oleh beragam peneliti sebelumnya (J. Jiang & Zhai, 2007)(Mihalkova et al., 2007)(Gao et al., 2008), yang kemudian melahirkan istilah *instance-transfer*, *feature-based transfer*, *parameter-transfer*, dan *knowledge-transfer*(Pan & Yang, 2010). Setiap pendekatan ini memiliki ciri khas masing-masing dalam proses melakukan transfer antar domain. Beragam pendekatan ini juga memiliki kesesuaian dengan kasus yang berbeda (Zhuang et al., 2019). Metode yang digunakan untuk setiap pendekatan ini juga beragam, mulai dari menghitung kesamaan, pencarian jarak, pembobotan, sampai dengan menggunakan struktur data lokal dari domain tersebut (Wang, Chen, et al., 2018; Wang et al., 2019; Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Zhong et al., 2018; Wang et al., 2017).

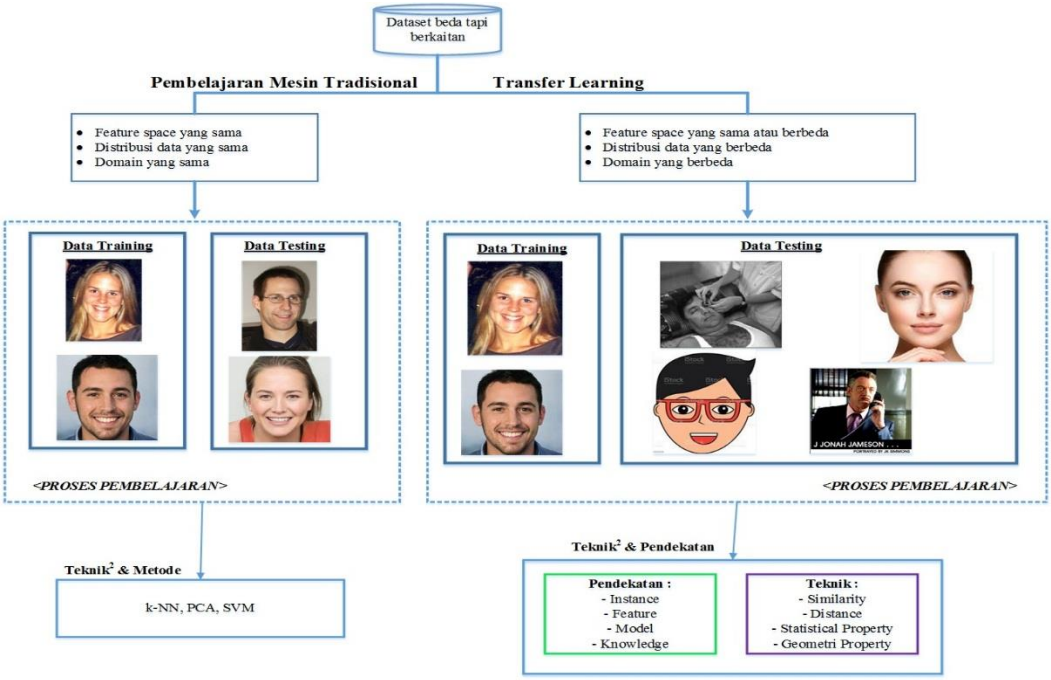
2.1 Perbedaan Pembelajaran Mesin dengan *Transfer Learning*

Pada pembelajaran mesin biasa, domain asal disebut sebagai data *training* dan domain tujuan atau target disebut sebagai data *testing* (Shao et al., 2015). Data *training* dan *testing* yang digunakan pada pembelajaran mesin juga masih berasal dari dataset yang sama, hanya dilakukan pembagian sehingga beberapa *instances* pada dataset akan digunakan sebagai data *training* dan sisanya digunakan sebagai data *testing*. Karena berasal dari dataset yang sama, baik data *training* dan *testing* akan memiliki *feature space* dan juga distribusi data yang sama.

Sebagai ilustrasi, maka Gambar 2.1 berikut ini menjelaskan tentang perbedaan antara pembelajaran mesin biasa dengan *transfer learning*.

Pada Gambar 2.1 dibawah, dapat dilihat bahwa pada *transfer learning* proses pembelajaran bisa melibatkan domain yang tidak sama persis. Gambar 2.1 mengilustrasikan bahwa ada 2 label yang akan diprediksi berdasarkan data yang diberikan menggunakan teknik/metode/pendekatan tertentu. Antara *transfer learning* dan pembelajaran mesin tradisional dapat menerapkan teknik/metode yang sama ataupun berbeda sama sekali, tergantung pendekatan yang digunakan pada *transfer learning*.

Perbedaan domain yang dimaksud pada *transfer learning* disini bisa terjadi karena perbedaan bentuk, perbedaan warna, perbedaan ukuran. Bahkan untuk domain gambar, perbedaan sudut pengambilan gambar juga bisa menyebabkan perbedaan distribusi data. Dua domain yang berbeda namun mempunyai kemiripan disebut sebagai *cross-domain*. Perbedaan tersebut bisa terletak pada distribusi *instances*, *instances space*, *feature space*, dan juga label *space* (Kouw & Loog, 2018; Kouw & Loog, 2021).



Gambar 2.1. Ilustrasi Perbedaan Pembelajaran Mesin Biasa dengan Transfer Learning

2.2 Cross-domain

Cross-domain adalah sebutan bagi dua domain yang berbeda namun masih saling berkaitan. Perbedaan kedua domain terletak pada perbedaan *feature space* dan distribusi data. Walaupun *transfer learning* dapat diterapkan pada domain yang memiliki perbedaan distribusi data, perlu dipastikan bahwa domain asal dan domain target saling berkaitan. Kaitan antar domain ini dapat diketahui dengan menghitung kemiripan antar domain. Perhitungan kemiripan ini dilakukan untuk menghindari terjadinya *negative transfer learning*. *Negative transfer* adalah kondisi yang mana antara domain asal dan domain target hanya memiliki sedikit kemiripan, sehingga proses *transfer learning* akan memberikan efek negatif pada domain target (Weiss et al., 2016a; Day & Khoshgoftaar, 2017; Zhuang et al., 2019). Terdapat banyak cara untuk mengukur kemiripan antar domain, seperti menggunakan *Maximum Mean Discrepancy* (MMD), *Kullback-Leibler Divergence*, *Bregman Divergence*, *Jensen-Shannon Divergence*(Zhuang et al., 2019). Dari beragam cara pengukuran kemiripan antar domain, MMD merupakan cara yang paling banyak digunakan pada beragam aplikasi *transfer learning* dan bekerja dengan menghitung perbedaan distribusi *marginal* antar domain asal dan domain target (Wang et al., 2019; Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Long, Wang, Ding, Pan, et al., 2014). MMD bekerja dengan menghitung perbedaan distribusi berdasarkan jarak rata-rata antara domain asal dan domain target (Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Long, Wang, Ding, Pan, et al., 2014).

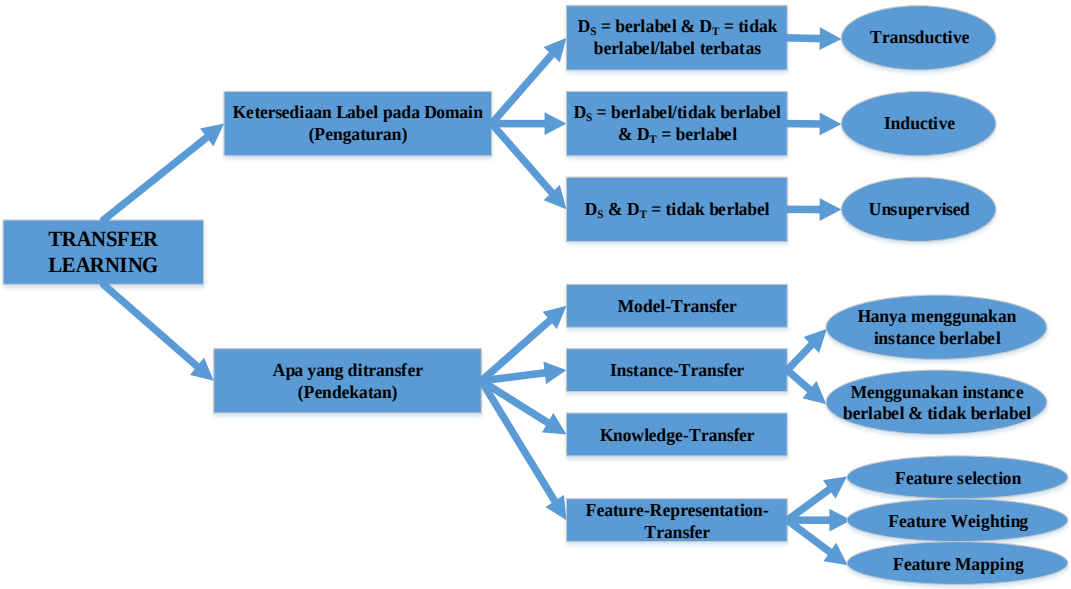
2.3 Pengelompokkan *Transfer Learning*

Seperti sudah disebutkan pada Bab 1 diatas, bahwa berdasarkan kesamaan *feature space*, *transfer learning* dapat dibagi menjadi dua, yaitu: *transfer learning* homogen dan *transfer learning* heterogen. *Transfer learning* homogen adalah kondisi pada saat terdapat persamaan *feature space* antara domain asal dan domain target. Beberapa contoh algoritma untuk *transfer learning* homogen adalah *Transfer Component Analysis* (TCA), *Joint Domain Adaptation* (JDA), *Adaptation Regularization based Transfer learning* (ARTL), *Geodesic Flow Kernel* (GFK), *Spectral Feature Allignment* (SFA), *2 Stage Weighting-Multi Source Domain Adaptation* (2SW-MDA), *Feature Augmented Method* (FAM), *Multi-Source TrAdaBoost* (MsTrAdaBoost), *Sample Selection and Feature Ensemble*(SSFE) (Weiss et al., 2016a; Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Long, Wang, Ding, Pan, et al., 2014; Long

et al., 2013a; Xia et al., 2015). Sementara *transfer learning* heterogen adalah *transfer learning* algoritma untuk kondisi adanya perbedaan *feature space* antara domain asal dan domain target. Contoh algoritma heterogen transfer antara lain adalah *Co-Regularized Online Transfer learning* (COTL), *Online Heterogenous Transfer with Weighted Classifier* (OHTWC), *Heterogenous Feature Augmentation* (HFA), *Domain Adaptation using Manifold Alignment* (DAMA), *Semi-Supervised Kernel Matching Domain Adaptation* (SSKMDA), *Correlation-Transfer SVM* (Day & Khoshgoftaar, 2017). Sementara berdasarkan ketersediaan *class* label baik pada domain asal maupun domain target, pengaturan *transfer learning* dibagi menjadi tiga kelompok, yaitu *inductive transfer learning*, *transductive transfer learning*, dan *unsupervisedtransfer learning*(Pan & Yang, 2010). Pada saat *class* label hanya terdapat pada domain asal, maka pengaturan yang digunakan adalah *transductive transfer*. Pada saat *class* label juga tersedia di domain target, maka pengaturan yang digunakan adalah *inductive transfer*. Sementara pada saat tidak tersedia *class* label pada domain asal dan domain target, maka pengaturan yang digunakan adalah *unsupervised transfer*. Pengaturan *inductive transfer learning* dapat meliputi *feature-representation transfer, instance-transfer, model-transfer*, dan *relational-knowledge-transfer*. Sementara untuk pengaturan *transductive transfer learning* bisa meliputi pendekatan *instance-transfer* dan *feature-representation transfer*. Untuk pengaturan *unsupervised transfer learning* hanya bisa meliputi *feature-representation transfer*.

Pada *instance-transfer* akan digunakan data berlabel pada domain asal yang sudah diberi bobot untuk membantu proses prediksi label di domain target (J. Jiang & Zhai, 2007). Pada *feature-representation transfer*, metode yang digunakan adalah mencari *feature representation* yang "baik" berdasarkan kemiripan yang diukur menggunakan *metric similarity* tertentu, seperti misalnya *Kullback-Leibler divergence* (Zhong et al., 2018). Untuk *model-transfer* metode yang dilakukan adalah mencoba untuk menemukan *shared parameter* dari model yang dibuat antara domain asal dan target (Mihalkova et al., 2007). Sementara pada *relational-knowledge-transfer* akan dibuat sebuah *mapping* hubungan antar pengetahuan antara domain asal dan target, misalkan dengan menggunakan *ensemble framework* yang membuat proses prediksi *class* label dapat sesuai dengan kondisi domain target tersebut (Gao et al., 2008). Gambar 2.2 menunjukkan pengelompokkan pengaturan dan pendekatan pada *transfer learning*. Sementara Tabel 2.1 menunjukkan perbedaan antara masing-masing pendekatan.

Berdasarkan ketersediaan label, selain *transfer learning*, maka dikenal juga istilah *zero-shot learning* dan *few-shot learning*. *Zero-shot learning* merupakan pendekatan *learning* yang mana tidak ada label pada domain asal. *Zero-shot learning* juga dapat dikatakan sebagai proses *learning* yang mana label domain asal dan label domain tujuan *disjoint* (Saad et al., 2022). Sementara pada *few-shot learning*, masih terdapat sedikit label pada domain asal (Xiaomeng et al., 2020).



Gambar 2.2. Pengelompokkan Transfer Learning

Kemampuan *transfer learning* untuk dapat melakukan melakukan prediksi secara efektif dan efisien antar domain yang berbeda membuat penelitian dan implementasi *transfer learning* berkembang dengan cukup cepat. Penelitian tentang *transfer learning* ini sendiri mulai muncul setelah tahun 2000 dan setiap tahun semakin beragam penelitian *transfer learning* yang dilakukan. Tabel 2.2 menjelaskan tentang beberapa penelitian tentang *transfer learning* termasuk jenis pendekatan yang digunakan pada penelitian tersebut.

Tabel 2.1. Pendekatan pada Transfer Learning

Pendekatan	Deskripsi
Instance-transfer	Menggunakan *instance* pada domain asal untuk digunakan dalam proses prediksi domain tujuan, misalkan dengan memberikan tambahan bobot pada *instance* yang akan digunakan pada proses prediksi.
Features-Representation-transfer	Menemukan dan menentukan *"good" features* antara domain asal dan domain tujuan. *"Good" features* ini dapat mengurangi perbedaan antar domain atau memperkecil *error* pada proses prediksi.
Model-transfer	Mencari dan menemukan paramater atau *prior* model yang dapat dibagikan antar domain.
Knowledge-transfer	Membuat *mapping knowledge* atau relasi antar domain.

Tabel 2.2. Penelitian-penelitian tentang Pendekatan Transfer Learning

No	Peneliti	Tahun	Pengaturan	Pendekatan
1	Blitzer et.al. (John Blitzer Ryan McDonald Fernando Pereira, 2006)	2006	*Transductive transfer learning*	*Feature-representation-transfer*
2	Jiang & Zhai (J. Jiang & Zhai, 2007)	2007	*Transductive transfer learning*	*Instance-transfer*
3	Mihalkova, Huynh, & Mooney (Mihalkova et al., 2007)	2007	*Transductive transfer learning*	*Knowledge-transfer*
4	Davis & Domingos (Davis & Domingos, 2009)	2008	*Transductive transfer learning*	*Knowledge-transfer*
5	Gao et.al. (Gao et al., 2008)	2008	*Transductive transfer learning*	*Model-transfer*
6	Pan, Kwok, & Yang (Pan et al., 2008a)	2008	*Transductive transfer learning*	*Feature-representation-transfer*
7	Pan et.al. (Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011)	2011	*Transductive transfer learning*	*Feature-representation-transfer*
8	Chattopadhyay et.al. (Chattopadhyay et al., 2012)	2011	*Transductive transfer learning*	*Instance-transfer & Model-transfer*

9	Long et.al. (Long et al., 2012)	2012	*Transductive transfer learning*	*Feature-representation-transfer*
10	Long et.al. (Long, Wang, Ding, Shen, et al., 2014)	2014	*Transductive transfer learning*	*Feature-representation-transfer*
11	Xia et.al. (Xia et al., 2015)	2015	*Transductive transfer learning*	*Instance-transfer & Feature-representation-transfer*
12	Wang et.al. (Wang et al., 2017)	2018	*Transductive transfer learning*	*Feature-representation-transfer*
13	Wang et.al. (Wang, Feng, et al., 2018a)	2018	*Transductive transfer learning*	*Feature-representation-transfer*
14	Cheplygina et.al. (Cheplygina et al., 2018)	2018	*Transductive transfer learning*	*Instance-transfer & Feature-representation-transfer*
15	Zhong et.al. (Zhong et al., 2018)	2018	*Transductive transfer learning*	*Feature-representation-transfer*
16	Wang et.al. (Wang et al., 2019)	2019	*Transductive transfer learning*	*Instance-transfer*
17	Han et.al. (Han et al., 2019)	2019	*Transductive transfer learning*	*Instance-transfer & Feature-representation-transfer*
18	Chen et.al. (Wang, Chen, et al., 2018)	2019	*Transductive transfer learning*	*Instance-transfer*
19	Jiang et.al. (S. Jiang et al., 2019)	2019	*Transductive transfer learning*	*Instance-transfer*

2.4 Komponen dalam *Transfer Learning*

Terdapat dua bagian dalam *transfer learning*, yaitu domain (D) dan *task* (T). Domain juga mengandung dua bagian, yaitu *feature space* $\mathcal{X}$ dan distribusi peluang marginal $P(X)$. Yang mana X adalah *instances set*, atau dituliskan $X = \{x_1, x_2, \dots, x_n\} \in \mathcal{X}$, sehingga sering dituliskan bahwa $D = \{\mathcal{X}, P(X)\}$. Sedangkan *task* juga mengandung dua bagian yaitu label

space $\mathcal{Y}$ dan fungsi prediksi $f(\cdot)$, dan sering dituliskan dengan $T = \{\mathcal{Y}, f(\cdot)\}$. Fungsi prediksi $f(\cdot)$ merupakan fungsi yang digunakan untuk memprediksi label *instance* pada domain target berdasarkan pembelajaran dari *feature vector* dan label *instance* pada domain asal. Bila dilihat dari sudut pandang statistik, maka $f(x) = P(y|x)$, atau dibaca peluang muncunya label y bila diberikan *instance* x (Pan & Yang, 2010; Weiss et al., 2016a; Zhuang et al., 2021).

Untuk domain asal (D_S), maka dapat dituliskan sebagai, $D_S = \{(x_{S1}, y_{S1}), (x_{S2}, y_{S2}), \dots, (x_{Sn}, y_{Sn})\}$, yang mana $x_{Si} \in \mathcal{X}_S$, adalah *instance* ke-i dari D_S dan $y_{Si} \in \mathcal{Y}_S$, adalah label yang terkait untuk *instance* x_{Si}. Demikian pula untuk domain tujuan atau target (D_T), yang dituliskan dengan $D_T = \{(x_{T1}, y_{T1}), (x_{T2}, y_{T2}), \dots, (x_{Tn}, y_{Tn})\}$, yang mana $x_{Ti} \in \mathcal{X}_T$, adalah *instance* ke-i dari D_T dan $y_{Ti} \in \mathcal{Y}_T$, adalah label yang terkait untuk *instance* x_{Ti} (Pan & Yang, 2010; Weiss et al., 2016b; Kouw & Loog, 2021; Kouw & Loog, 2018).

2.5 Beragam Teknik Pembelajaran pada *Transfer Learning*

Proses pembelajaran pada *transfer learning* juga bisa menggunakan algoritma dan teknik yang sama atau berbeda dengan yang digunakan pada pembelajaran mesin biasa. Setiap pendekatan yang digunakan pada penelitian *transfer learning* menggunakan beragam algoritma dan teknik berbeda-beda untuk mengatasi perbedaan distribusi data maupun *feature* antar domain. Tabel 2.3 berikut merangkumkan beberapa penelitian terbaru yang menggunakan pendekatan *instance-transfer*, *feature-representation-transfer*, atau pendekatan gabungan beserta dengan beragam teknik yang digunakan.

Tabel 2.3. Penelitian tentang Beragam Teknik pada Transfer Learning

No	Peneliti	Tahun	Algoritma/Teknik/Metode			
			Kemiripan	Bobot	Struktur Data Lokal	Teknik Lainnya
1	Long et.al. (Long, Wang, Ding, Pan, et al., 2014)	2014	√ (MMD)		√ (Rata-rata)	
2	Xia et.al.(Xia et al., 2015)	2015	√ (KL)	√ (*instance*)		
3	Wang et.al.(Wang et al., 2017)	2018	√ (MMD)	√ (*class label*)		
4	Wang et.al.(Wang, Feng, et al., 2018a)	2018	√ (MMD)		√	

5	Cheplygina et.al.(Cheplygina et al., 2018)	2018		√ (*instance*)		
6	Zhong et.al.(Zhong et al., 2018)	2018	√ (KL)	√ (*feature*)		
7	Wang et.al.(Wang et al., 2020)	2019	√ (MMD)		√ (Rata-rata)	
8	Han et.al.(Han et al., 2019)	2019	√ (Cosine)			
9	Chen et.al.(Wang, Chen, et al., 2018)	2019	√ (MMD)			√ (Jarak)
10	Jiang et.al.(J. Jiang & Zhai, 2007)	2019	√ (MMD)	√ (*instance*)	√ (Properti Geometrik)	

Penggunaan jarak pada penelitian sebelumnya tentang *transfer learning* pada umumnya dilakukan untuk menghitung kemiripan atau sebagai fungsi *mapping*. Penjelasan (Sharma & Batra, 2014) untuk beberapa rumus jarak yang umum digunakan pada kasus gambar adalah:

- ***Euclidean Distance***, merupakan perhitungan jarak yang umum digunakan untuk menghitung kemiripan antar *feature vector*. *Euclidean Distance* juga umum digunakan pada penelitian *feature-representation transfer learning* pada fungsi *mapping* yang digunakan (Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Long et al., 2013b; Long, Wang, Ding, Sun, et al., 2014; Tahmoresnezhad & Hashemi, 2017b; Wang, Feng, et al., 2018b; Wang et al., 2017; Wang et al., 2019). Persamaan *Euclidean Distance* dapat dilihat pada Persamaan 2.1 berikut:

$$ED(x, y) = \sqrt{\sum_{i=1}^{N}(x_i - y_i)^2}, \quad (2.1)$$

- ***Manhattan Distance***, merupakan perhitungan jarak menggunakan garis lurus, baik horisontal maupun vertikal dan juga digunakan untuk menghitung kemiripan antar *feature vector*. Persamaan *Manhattan Distance* dapat dilihat pada Persamaan 2.2 berikut:

$$MD(x, y) = \sum_{i=1}^{N}|x_i - y_i|, \quad (2.2)$$

- ***Mahalanobis Distance***, merupakan perhitungan jarak yang dipengaruhi juga oleh nilai *covariance* matrik dengan bentuk persamaan dapat dilihat pada Persamaan 2.3.

$$d(x, y) = \sqrt{(x-y)^T S^{-1}(x-y)}, \quad (2.3)$$

yang mana ED, MD, d adalah *Euclidean Distance*, *Manhattan Distance*, dan *Mahalanobis Distance* dengan x, y adalah *point* yang akan dihitung jaraknya. *Point* ini dapat berupa vector, dengan x_i berarti bahwa point x yang ke-i. N adalah jumlah total *instances* atau total *point* yang akan dihitung jaraknya. S adalah *covariance* matrik.

2.6 Feature-Representation Transfer (Feature-based Transfer Learning)

Pada Gambar 2.2 dapat dilihat bahwa berdasarkan pada apa yang ditransfer, *transfer learning* dapat dibagi menjadi *instance-transfer learning*, *feature-representation transfer*, *model-transfer*, dan *relational-knowledge-transfer*. Perbedaan karakteristik kerja umum dari setiap kategori tersebut dituliskan pada Tabel 2.1. Dari Tabel 2.1 juga dapat dilihat bahwa *feature-representation transfer* bekerja dengan mencari representasi *feature* yang "baik" pada *cross-domain*. Representasi *feature* yang "baik" ini adalah representasi yang masih dapat mempertahankan struktur data lokal yang penting dari masing-masing domain (Zhuang et al., 2021).

(Zhuang et al., 2021) menyebutkan bahwa *feature-representation transfer* merupakan kelompok *data-based interpretation*, yang berarti fokus pada proses penyelarasaan *cross-domain* dengan cara melakukan transformasi data. *Data-based interpretation* sendiri dibagi menjadi dua kategori *transfer learning*, yaitu *instance-transfer* dan *feature-representation transfer*. Pada *feature-representation transfer*, transformasi data yang dimaksud adalah transformasi *feature*.

Transformasi *feature* dilakukan untuk mengurangi perbedaan antar domain dengan cara membentuk representasi *feature* yang baru. Terdapat beberapa metrik yang diadopsi untuk mengurangi perbedaan antar domain, antara lain: *Maximum Mean Discrepancy (MMD), Kullback-Leibler Divergence, Bregman Divergence, dan Jensen-Shannon Divergence*. Diantara metrik- metrik ini, MMD merupakan metirk yang paling banyak digunakan untuk *transfer learning* pada kasus klasifikasi.

Transformasi *feature* dilakukan juga dapat dilakukan dengan menerapkan beberapa strategi, seperti :

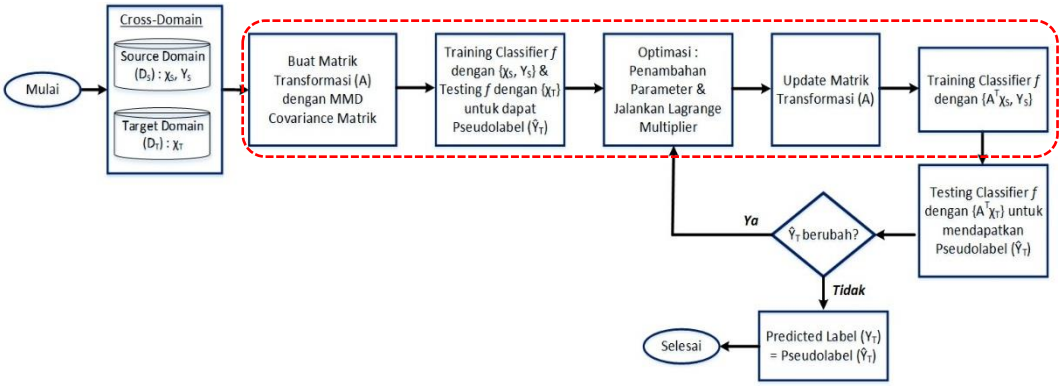
- ***Feature Augmentation***, biasa diterapkan dengan cara melakukan replikasi *feature* atau *stacking feature*, untuk mendapatkan *feature space* yang bersifat latent berdasarkan "struktur yang penting".
- ***Feature Mapping***, biasa diterapkan menggunakan fungsi *mapping* untuk mengurangi perbedaan distribusi antar domain dengan tetap mempertahankan struktur data.
- ***Feature Clustering***, bertujuan untuk menemukan representasi *feature* yang abstrak dari *feature* asli dan biasa diterapkan untuk kasus klasifikasi dokumen.
- ***Feature Selection***, biasa bekerja dengan cara melakukan ekstraksi terhadap *pivot features*, atau *features* yang mempunyai *behavior* sama pada *cross-domain*.
- ***Feature Encoding***, bekerja menggunakan prinsip *autoencoder* yang akan melakukan *reshaping* terhadap *features*, sehingga banyak diadopsi pada area *deep learning*.
- ***Feature Alignment***, yang akan fokus pada beberapa *features* implisit dan bekerja dengan pendekatan *subspace alignment*.

Beberapa penelitian *feature-representation transfer learning* sebelumnya juga fokus pada upaya untuk mendapatkan *transformation matrix* atau matrik adaptasi. Matrik adaptasi ini terbentuk dengan menerapkan kernel pada rumus MMD yang digunakan untuk membentuk representasi *feature* baru dengan cara pengurangan dimensi pada dataset yang digunakan (Pan et al., 2008b; Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Long et al., 2013b; Wang, Feng, et al., 2018b;Wang et al., 2017;Gong et al., 2012; Tahmoresnezhad & Hashemi, 2017a). Berdasarkan penelitian sebelumnya tentang *feature-representation transfer learning* juga menunjukkan adanya fokus pada masalah optimasi untuk meningkatkan unjuk kerja metode dengan menambahkan beberapa parameter optimasi sesuai dengan tujuan dan karakteristik masing-masing metode (Wang, Feng, et al., 2018b;Wang et al., 2017;Tahmoresnezhad & Hashemi, 2017a; Wang et al., 2020; Long, Wang, Ding, Sun, et al., 2014).

Berdasarkan proses transformasi *feature* dan cara optimasi metode yang dilakukan untuk menghasilkan label hasil prediksi, *feature-based transfer learning* dibagi menjadi dua kelompok, yaitu kelompok yang bekerja secara iterasi dan kelompok yang bekerja tanpa iterasi (*easy*). Berikut adalah penjelasan dari masing-masing kelompok *feature-represesentation transfer learning* tersebut.

A. Iterative Feature-based Transfer Learning

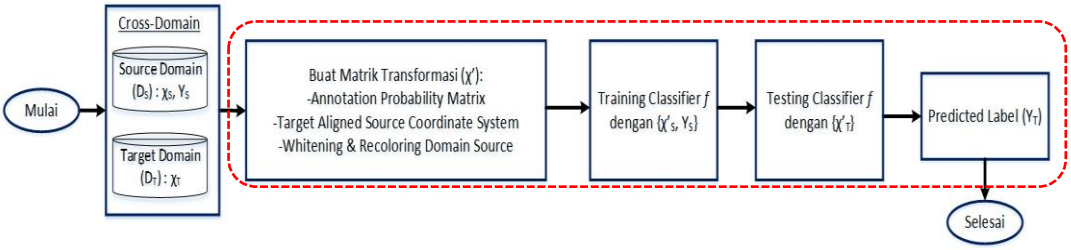
Berdasarkan penelitian sebelumnya dapat dilihat ada dua pendekatan yang umum digunakan. Pendekatan umum pertama adalah dengan menggunakan pendekatan *dimensionality reduction* dan melakukan proses iterasi untuk mendapatkan *pseudolabel* ($\widehat{\mathcal{Y}_T}$) bagi data tidak berlabel pada domain target, atau *unlabaled instances* di D_T. Proses iterasi ini baru berhenti pada saat $\widehat{\mathcal{Y}_T}$ sudah *converged*, atau tidak berubah lagi nilainya setelah beberapa iterasi berturut-turut. Iterasi perlu dilakukan karena pada awal proses prediksi akan dihasilkan $\widehat{\mathcal{Y}_T}$ yang tidak tepat, mengingat belum dilakukan pembelajaran pada *cross-domain*. Pendekatan *dimensionality reduction* pada pendekatan umum pertama ini, dilakukan sebagai bentuk proses pembelajaran dengan membentuk matrik adaptasi atau disebut MMD *covariance matrix* yang dibentuk dari penggabungan *instances* pada *cross-domain* menggunakan rumus MMD dan kernel. Tahap ini merupakan tahap untuk mengurangi perbedaan distribusi marginal dan conditional pada *cross-domain*. Sementara optimasi akan dilakukan pada fungsi pembelajaran (*learning function*) yang mengandung sejumlah parameter regularisasi dan dilakukan untuk mendapatkan matrix transformasi orthogonal. Matrix transformasi orthogonal ini yang kemudian akan digunakan untuk *training* dan *testing* sehingga diharapkan mendapatkan $\widehat{\mathcal{Y}_T}$ yang lebih baik. Sejumlah metode yang menggunakan pendekatan pertama ini antara lain adalah TCA, JDA, TJM, BDA, MEDA, VDA (Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, 2011; Long et al., 2013b; Long, Wang, Ding, Sun, et al., 2014; Wang et al., 2017; Wang, Feng, et al., 2018a; Tahmoresnezhad & Hashemi, 2017a). Tahapan umum dari pendekatan *iterative feature-based transfer learning* ini dapat dilihat pada Gambar 2.3 berikut ini. Bagian kotak merah menunjukan proses pembelajaran atau proses transformasi *feature* yang dilakukan.



Gambar 2.3. Tahapan Umum Pendekatan dengan Iterasi pada Feature-based Transfer Learning

B. *Easy Feature-based Transfer Learning*

Pendekatan umum kedua pada *feature-based transfer learning*, yang merupakan *easy feature-based transfer learning* diusung oleh metode *Easy Transfer Learning* (Wang et al., 2019), *correlation alignment* (Coral) (Sun et al., 2016), dan *subspace alignment* (SA) (Fernando et al., 2014). Berbeda dari pendekatan pertama, pada pendekatan kedua ini, tidak dilakukan iterasi untuk mendapatkan $\widehat{\mathcal{Y}_T}$, karena proses *training* dan *testing* dilakukan setelah proses pembelajaran pada *cross-domain*, untuk membentuk matrik transformasi. Pada pendekatan kedua ini juga tidak digunakan rumus MMD dan kernel sebagai upaya mengurangi perbedaan distribusi pada *cross-domain*. Sebagai contoh pada *Easy Transfer Learning*, untuk mengurangi perbedaan distribusi dilakukan dengan memanfaatkan *unlabaled instances* pada D_T, melalui perhitungan jarak *instances* ke setiap pusat *class label*. Cara ini disebut dengan membentuk *probability annotation matrix*. Label hasil prediksi atau *predicted label* ($\mathcal{Y}_T$) dari suatu *instances* didapat dengan mengambil indek *class label* yang mempunyai jarak terdekat dengan *instances* tersebut. Sementara pada *correlation alignment*, pengurangan distribusi dilakukan dengan melakukan *subspace alignment* pada *cross-domain* dengan cara melakukan *re-coloring* pada D_T sehingga menjadi lebih mirip dengan D_S. Pendekatan kedua ini lebih sederhana dalam implementasinya, tidak memerlukan *resource* besar untuk menjalankan proses pembelajarannya, sehingga memungkinkan untuk implementasi pada kondisi keterbatasan *resource*. Namun pada pendekatan kedua ini selain kurang memperhatikan kesamaan antar domain sehingga memungkinkan terjadinya *negative transfer learning*, juga dalam proses penyamaan distribusi lebih banyak menggunakan informasi pada D_S. Tahapan umum dari kelompok *easy feature-based transfer learning* ini dapat dilihat pada Gambar 2.4, yang mana kotak merah memperlihatkan proses pembelajaran dan transformasi *feature* yang dilakukan oleh *metode feature-transfer learning* yang menggunakan pendekatan *easy feature-representation transfer*.



Gambar 2.4. Tahapan Umum Pendekatan tanpa Iterasi pada Feature-based Transfer Learning

BAB 3
METODOLOGI PENELITIAN

3.1. Gambaran Umum WbFTL

Penelitian ini mengusung *transductive transfer learning*, yang mana *instance* berlabel hanya dimiliki oleh domain asal. Domain tujuan mengandung *instance* tidak berlabel atau hanya mengandung sedikit saja *instance* berlabel (*limited labeled instance*). Pada prinsipnya metode yang diusulkan ini dapat digunakan untuk *supervised* dan *semi-supervised*. Tugas dari kedua domain ini juga sama, yaitu klasifikasi. Baik domain asal dan domain tujuan merupakan dataset gambar dengan jumlah *feature* sama ($bF_S = bF_T$). Sementara jumlah *instances* pada kedua domain bisa berbeda ($n_S \neq n_T$).

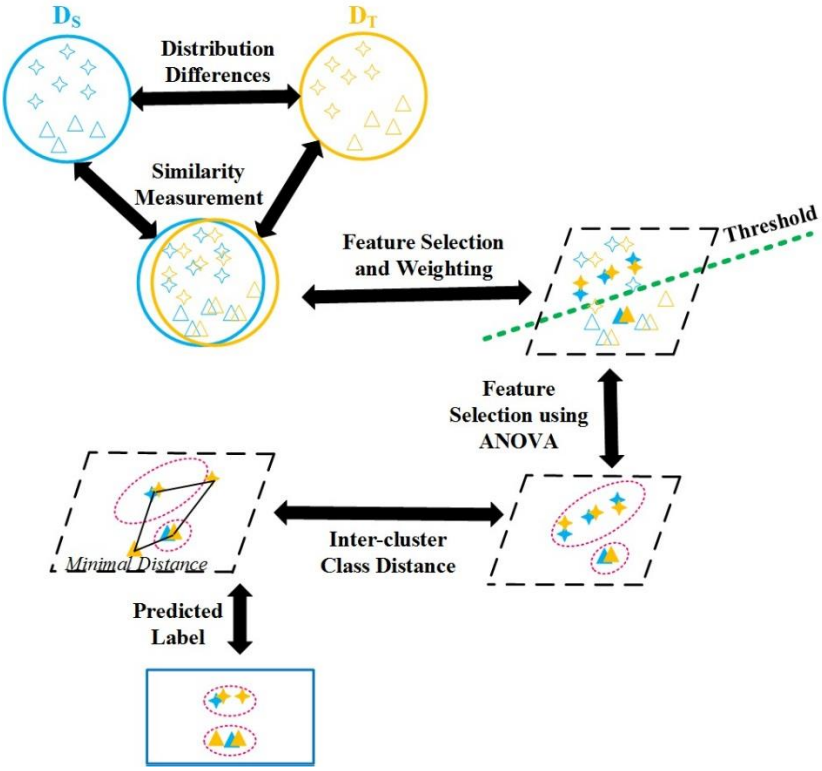
Kedua domain yang digunakan berasal dari dataset yang berbeda namun berkaitan, sehingga mengandung perbedaan distribusi, baik marginal maupun *conditional distribution*. Oleh karena itu perlu dilakukan *distribution adaptation* pada *cross-domain* ini. *Distribution adaptation* merupakan kasus khusus dalam *transfer learning*, yang bertujuan untuk meningkatkan unjuk kerja model pada saat menangani *cross-domain* dengan perbedaan distribusi namun task yang sama. Sementara *transfer learning* adalah kasus khusus dari pemelajaran mesin yang dapat menangani kondisi yang mana task pada *cross-domain* juga berbeda.

WbFTL menggunakan pendekatan sederhana (*easy*) dengan melakukan pengembangan serta beberapa modifikasi untuk mengurangi perbedaan distribusi pada *cross-domain* dan upaya peningkatan hasil akurasi. Istilah pendekatan sederhana atau *easy* hanya akan digunakan pada konteks penelitian ini untuk membedakan dengan pengelompokkan lain, yaitu iteratif. Istilah sederhana atau *easy* mengarah pada penggunaan minimal parameter dan tanpa perlu iterasi berkali-kali untuk mendapatkan label hasil prediksi. Sehingga dapat menghemat sumber daya karena tidak perlu memproses banyak parameter.

A. Strategi Transformasi *Feature* yang Digunakan

Pengembangan yang dilakukan terletak pada strategi transformasi *feature* yang dilakukan dengan metode pemilihan *feature* dan pemberian bobot kepada *feature* terpilih. Penggunaan metode pemilihan *feature* dan pemberian bobot ini membuat WbFTL merupakan *non-parametric method*, karena tidak membutuhkan banyak parameter dalam proses pembelajaran

yang dilakukan. Sementara modifikasi dilakukan pada cara untuk mengatasi perbedaan distribusi pada *cross-domain*. Pada metode WFbTL ini, untuk mengatasi perbedaan distribusi antar domain dilakukan dengan menerapkan rumus MMD dan kernel untuk membentuk representasi *feature* yang baru dari *cross-domain*. Pemanfaatan *unlabaled instances* pada domain tujuan melalui perhitungan jarak ke setiap class label dilakukan untuk mengatasi perbedaan *conditional distribution*. Selain itu, dalam rangka mengatasi perbedaan *conditional distribution* juga dilakukan dengan menerapkan ANOVA (*Analysis of Variance*) untuk memilih hanya *feature* yang signifikan antar class label dan dalam class label. *Feature* yang terpilih inilah yang akan digunakan untuk proses training dan inference *classifier*, sehingga menghasilkakn predicted label (Y_T) bagi D_T. Gambaran umum dari WbFTL dapat dilihat pada Gambar 3.1 berikut.



Gambar 3.1. Gambaran Umum WbFTL

B. Penggunaan Properti Statistik

WbFTL menggunakan properti statistik yaitu rata-rata dan varians dalam setiap tahapan transformasi yang dilakukan. Mulai dari *feature* transformasi pertama yang menggunakan

rata-rata setiap *feature* sebagai fungsi mapping. Dalam pemilihan *feature* pada transformasi pertama ini juga menggunakan rata-rata sebagai *threshold* atau batas untuk menentukan *feature* yang terpilih. Demikian pula pada *feature* transformasi kedua yang menggunakan varians untuk memilih *feature* yang signifikan dalam class label dan antar class label. Perhitungan jarak pada *feature* transformasi ketiga juga menggunakan properti statistik, yaitu rata-rata. Rata-rata dan varians sendiri merupakan properti statistik yang sudah pasti dimiliki oleh setiap jenis data dan dapat memberikan gambaran distribusi data sehingga dapat digunakan untuk melakukan perbandingan antar kelompok (Frederick J Gravetter and Larry B. Wallnau, 2013). Penelitian sebelumnya juga mengungkapkan bahwa penggunaan properti statistik dapat menghemat resource pemrosesan data sehingga memungkinkan penerapan *transfer learning* pada *wearable device* (Wang et al., 2019).

3.1 Tahapan Pemelajaran

Pada dasarnya, terdapat empat tahapan pembelajaran, yang mana tahap pertama merupakan tahap untuk mengukur kesamaan antar domain. Tiga tahap setelahnya merupakan tahap transformasi *feature* pada WbFTL. Secara umum tahapan pembelajaran dan *feature* transformasi pada WbFTL dapat dituliskan sebagai berikut:

A. Tahap 1 : Tahap mengukur kesamaan antar domain.

Tujuan tahap ini adalah untuk mengurangi perbedaan distribusi marginal pada *cross-domain* dengan menerapkan MMD dan kernel untuk mengukur perbedaan antar domain. Hasil dari proses pembelajaran tahap 1 ini adalah sebuah vektor representasi *feature* berisi *feature mean* dari *cross-domain* hasil implementasi MMD dan kernel.

B. Tahap 2 : *Feature* Transformasi 1

Feature transformasi ini ini bertujuan untuk meningkatkan akurasi dengan memanfaatkan *feature* yang signifikan untuk klasifikasi. Transformasi pertama ini dilakukan dengan menerapkan teknik pemilihan *feature* dan pemberian bobot. Pemilihan *feature* pada transformasi pertama ini mengadopsi metode filter, yaitu memilih *feature* berdasarkan batas (*threshold*), yang mana *threshold* ditentukan berdasarkan nilai rata-rata dari vektor representasi *feature* hasil pembelajaran tahap 1. *Feature* yang terpilih kemudian akan diberikan bobot. Hasil dari *feature* transformasi 1 ini berupa vektor representasi *feature* yang dituliskan dengan notasi F_{Trans1}.

C. Tahap 3 : *Feature* Transformasi 2

Tahap 3 ini masih bertujuan untuk meningkatkan akurasi dengan teknik pemilihan *feature*. Teknik pemilihan *feature* yang digunakan adalah teknik ANOVA. Dengan menerapkan teknik ANOVA, akan diambil hanya *feature* yang signifikan terhadap label dan *feature* yang signifikan dalam label. Hasil *feature* transformasi 1 akan diproses dengan ANOVA dan menghasilkan vektor representasi *feature* yang dituliskan dengan notasi F_{Trans2}.

D. Tahap 4 : *Feature* Transformasi 3

Feature transformasi 3 ini hanya diterapkan pada D_T. Tahap ini bertujuan mengurangi perbedaan *conditional distribution* pada *cross-domain* dengan memanfaatkan *instances* tidak berlabel pada D_T. Transformasi *feature* 3 ini dilakukan dengan menghitung jarak *instances* ke pusat masing-masing *class* label (*inter-cluster class label*).

Gambaran proses *training* dan inferensi pada WbFTL dapat dilihat pada Gambar 3.2. Sementara tahapan *feature* transformasi yang dilakukan pada WbFTL dapat dilihat pada Gambar 3.3.

3.2. *Feature* Transformasi 1 dengan Pemilihan *Feature* (*Feature* Selection) menggunakan *Threshold*

Proses transformasi *feature* 1 pada WbFTL ini mengadopsi metode filter, yaitu memilih *feature* yang signifikan berdasarkan *threshold*. *Threshold* yang digunakan adalah rata-rata, karena alasan fleksibilitas dan scalabilitas. Rata-rata sudah pasti dimiliki oleh semua jenis data dan menggambarkan *central tendency* dari dataset yang digunakan. Alasan ini membuat *theshold* yang digunakan juga bersifat dinamis, karena akan langsung berubah menyesuaikan kondisi data. Penggunaan rata-rata sebagai *threshold* dalam pemilihan *feature* juga sudah digunakan pada kasus pemilihan *feature* (Güneş et al., 2010; Sevani et al., 2019).

*P*persamaan 3.1 merupakan persamaan untuk menghitung *threshold*. F_m ini sudah merupakan representasi *feature* baru yang menggambarkan *feature* dari domain asal dan domain tujuan.

$$Th = \frac{1}{n}\sum_{i=1}^{n} f_{mi} \tag{3.1}$$

yang mana nilai f_{mi} berasal dari $F_m = \{f_{m1}, f_{m2}, \dots, f_{mb}\} \in R^b$.

Proses pemilihan *feature*s menggunakan Th akan menghasilkan subset *feature* yang akan digunakan untuk proses berikutnya, seperti terlihat pada persamaan 3.2. Pada persamaan 3.2 ini merupakan tahapan pemilihan *feature* , yang mana F_z didapat dari F_m yang nilainya lebih kecil daripada *threshold* Th. F_z adalah vektor representasi *feature* yang terbentuk dengan cara

memilih *feature-feature* pada F_m yang nilainya kurang dari Th, seperti ditulis pada persamaan (3.3)

$$F_z = F_m < Th, \tag{3.2}$$

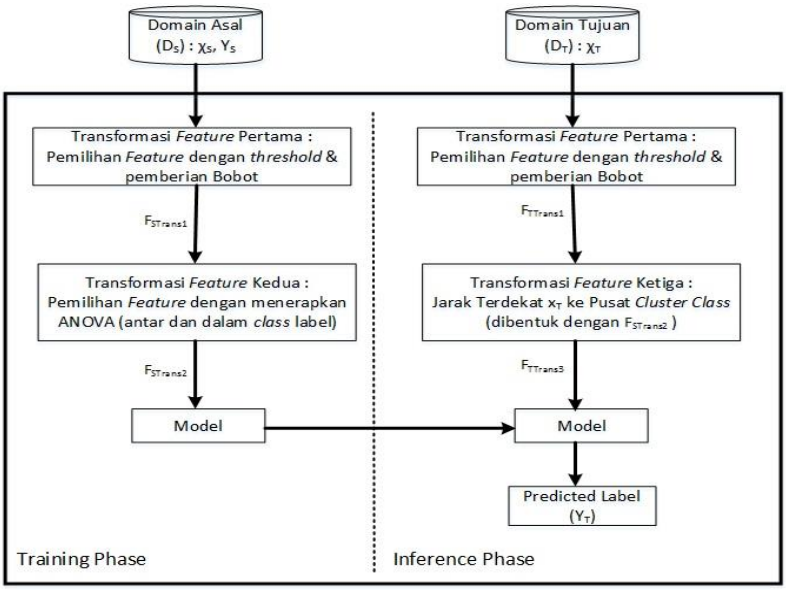
$$F_z = \{f_{z1}, f_{z2}, \dots, f_{zk}\}, \text{yang mana } k < b \tag{3.3}$$

Hasil *feature* terpilih pada langkah pertama ini yang kemudian akan diberikan bobot lokal, yang artinya setiap *feature* dapat mempunyai nilai bobot yang berbeda. Setelah seluruh *feature* pada F_z diberikan bobot, maka akan terbentuklah *feature* transformasi 1 (F_{Trans1}) untuk *instances* pada D_S dan D_T, seperti pada persamaan 3.5 dan persamaan 3.6.

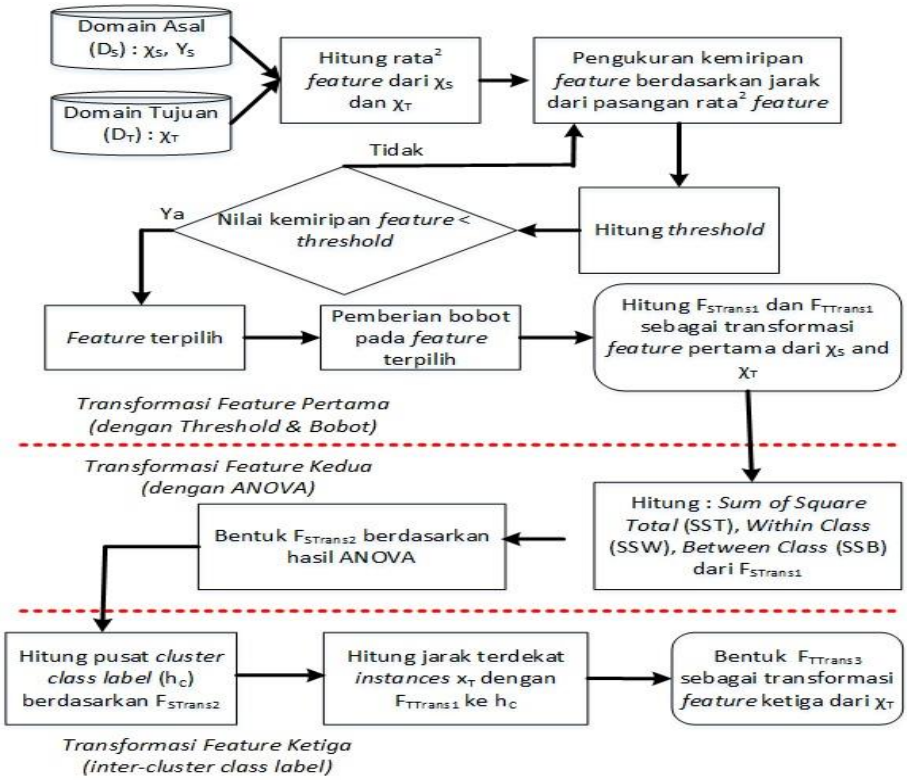
$$w(f_{zi}) = \frac{\frac{1}{k}\sum_{i=1}^{k} f_{zi}}{f_{zi}}, \tag{3.4}$$

$$F_{STrans1(i)} = w(f_{zi}) \times f_{Si}, \tag{3.5}$$

$$F_{TTrans1(i)} = w(f_{zi}) \times f_{Ti}, \tag{3.6}$$



Gambar 3.2. Proses Training dan Inference pada WbFTL



Gambar 3.3. Transformasi Feature pada WbFTL

3.3. Transformasi *Feature* 2 dengan Teknik Pemilihan *Feature* Menggunakan ANOVA

Dalam cara kerjanya WbFTL menerapkan dua teknik pemilihan *feature*, yang mana teknik pertama mengadopsi cara kerja metode filter yang menggunakan batas (*threshold*) untuk memilih *feature*. Teknik kedua adalah menerapkan ANOVA untuk memilih *feature* yang signifikan terhadap label. Pada WbFTL, pemilihan *feature* yang signifikan akan diterapkan pada D_S, mengingat bahwa ANOVA membutuhkan informasi label dan D_T tidak berlabel.

Konsep formal dari transformasi *feature* menggunakan ANOVA untuk proses pemilihan *feature* dapat dilihat pada persamaan 3.7 dan persamaan 3.8. Dimulai dari $F_{STrans1}$ pada persamaan 3.5, yang kemudian akan diukur dan dipilih *feature* yang signifikan antar label dan dalam label, sehingga membentuk $F_{STrans2}$. Nilai *feature* pada $F_{STrans1}$ sama dengan nilai

feature pada $F_{STrans2}$, hanya jumlah *feature* saja yang berbeda antara kedua subset *feature* tersebut.

$$F_{STrans2(i)} = F_{STrans1(i),} \tag{3.7}$$

$$F_{STrans2} = \{F_{STrans2(1),}F_{STrans2(2),} \cdots, F_{STrans2(a),}\} \in R^{a}, \tag{3.8}$$

yang mana $a < k < b$

3.4. Transformasi *Feature* 3 dengan *Inter-Cluster Class* Label

Penggunaan informasi label pada transformasi *feature* dengan ANOVA masih hanya melibatkan label dari D_S. Hal ini dirasa kurang menggambarkan label pada D_T. Mengingat bahwa yang akan diprediksi adalah label pada D_T, maka fokus pada informasi yang terkandung pada D_T diharapkan dapat memberikan hasil yang lebih optimal. Untuk dapat memanfaatkan informasi yang terkandung pada $D_{T,}$ maka dilakukan transformasi *feature* dengan menggunakan informasi jarak minimal dari *instances* pada D_T ke pusat (*centroid*) dari label-label yang ada. Penggunaan jarak juga melahirkan konsep klasifikasi yang mana data akan cenderung berkelompok dalam satu label yang sama bila mempunyai kesamaan dan akan cenderung berjauhan bila tidak mempunyai kesamaan. Kesamaan dan perbedaan data ini ditentukan oleh jarak antar data dan *features* yang terkandung dalam data.

Penggunaan jarak dan *centroid* label dilatar belakangi oleh cara kerja k-*Nearest Neighbor* (k-NN) dan juga hukum gravitasi Newton. Berdasarkan k-NN dapat dilihat bahwa penentuan label bagi suatu data dapat dilakukan dengan mengukur jarak dan mengambil label yang mempunyai jarak terdekat dengan data tersebut. Menggunakan konsep cara kerja hukum gravitasi juga dapat dilihat bahwa kekuatan daya tarik dua obyek dipengaruhi jarak dan massa dari obyek.

Transformasi *feature* dengan menggunakan jarak minimal ke pusat label akan menjadi tahap akhir proses transformasi, yaitu transformasi ketiga, sekaligus penentu hasil prediksi label untuk D_T. Dimulai dari persamaan 3.6 yang merupakan transformasi *feature* pertama pada D_T ($F_{TTrans1}$), maka nilai transformasi *feature* tersebut yang akan diterapkan pada *instances* D_T untuk membentuk $F_{TTrans3}$ dengan mengadopsi persamaan hukum gravitasi seperti pada persamaan 3.9 dan 3.10.

$$F_{TTrans3(i)} = \frac{F_{TTrans1(i)} \times h_C}{Dist\left(F_{TTrans1(i)}, h_C\right)^2}, \quad (3.9)$$

$$F_{TTrans3} = \left\{F_{TTrans3(1),} F_{STrans2(2),} \cdots, F_{STrans2(k),}\right\} \in R^k, \quad (3.10)$$

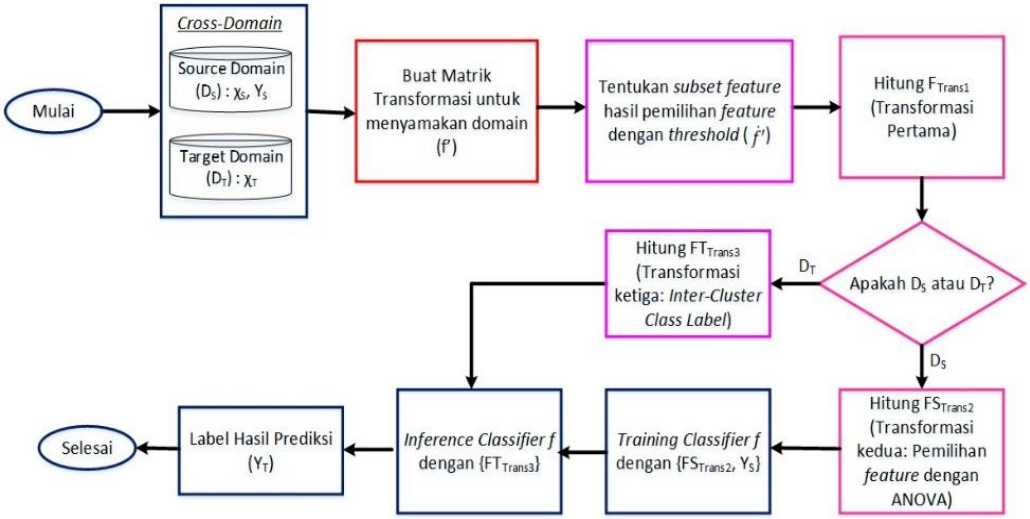
3.5. Resume Cara Kerja WbFTL

Tabel 3.1 menunjukkan resume perbedaan antara WbFTL dengan kedua kelompok *feature-based* transfer learning yang ada. Sementara Gambar 3.4 menggambarkan perbedaan WbFTL dengan metode *feature-transfer learning* lainnya.

Tabel 3.1. Resume Perbandingan WbFTL dengan Metode Feature-based Transfer Learning Lainnya

No	Prosedur	WbFTL	Iterative *Feature-based Transfer Learning* (Pendekatan Iterasi)	*Easy Feature-based Transfer Learning* (Pendekatan Langsung)
1	Input dan Output	Input : D_S yang berlabel dan D_T yang tidak berlabel. Output : *Classifier* dan $\mathcal{Y}_T$.	Input : D_S yang berlabel dan D_T yang tidak berlabel. Output : *Classifier* dan $\mathcal{Y}_T$.	Input : D_S yang berlabel dan D_T yang tidak berlabel. Output : *Classifier* dan $\mathcal{Y}_T$.
2	Perhitungan kemiripan antar domain	Menggunakan MMD seperti pada persamaan 3.5	Menggunakan MMD, mengikuti (Pan et al., 2008a)	Tidak melakukan perhitungan kemiripan antar domain.
3	Transformasi *feature*	Menggunakan pendekatan pemilihan *feature* dan menerapkan transformasi yang berbeda antara D_S dan D_T.	Menggunakan pendekatan dimensionality reduction, dengan membentuk matrik proyeksi dan optimalisasi dilakukan dengan Lagrange Multiplier, dan	Dilakukan dengan menyamakan *feature* D_S agar makin mirip dengan $D_{T,}$ melalui proses recoloring dan whitening atau perhitungan jarak *instances* ke cluster class label.

			dilakukan secara berulang. Proses transformasi yang sama diterapkan pada D_S dan D_T.	Proses transformasi yang sama diterapkan pada D_S dan D_T.
4	*Classifier*	SVM	1NN	SVM
5	Pseudolabel	Tidak menggunakan pseudolabel dalam proses training dan inference model untuk menghasilkan $\mathcal{Y}_T$.	Menggunakan pseudolabel dalam proses training dan inference model untuk menghasilkan $\mathcal{Y}_T$.	Tidak menggunakan pseudolabel dalam proses training dan inference model untuk menghasilkan $\mathcal{Y}_T$.
6	Hyperparameter setting	Tidak ada hyperparameter setting, kecuali pada parameter *classifier*.	Penambahan parameter regularisasi dan beberapa parameter lainnya, tergantung pada tujuan perancangan metode	Tidak ada hyperparameter setting, kecuali pada parameter *classifier*



Gambar 3.4. Ilustrasi Perbedaan Menyeluruh antara WbFTL dengan Metode Feature-based Transfer Learning Lainnya

BAB 4
HASIL PENELITIAN

4.1. Dataset dan Eksperimen Setting

Terdapat dua sumber awal untuk dataset yang digunakan pada penelitian ini, yaitu *Reseach Data Repository* dari *California Institute of Technology* dan *Electrical Engineering* dari UC Berkeley. Dataset yang digunakan pada penelitian ini kemudian mengalami *pre-processing* dengan menggunakan SURF. SURF sendiri merupakan algoritma *feature detector* dan *feature descriptor* yang dapat bekerja lebih efisien (cepat dan hemat memori) dibandingkan algoritma *feature detector* dan *descriptor* sejenis, seperti SIFT dan DAISY, dengan hasil yang tetap komparatif. Label yang digunakan pada penelitian ini juga dibatasi pada 10 label yang juga digunakan pada penelitian *feature-based* transfer learning sebelumnya. Tabel 4.1 menjelaskan tentang dataset yang digunakan. Gambar 4.1 menunjukkan contoh gambar asli yang digunakan.

Tabel 4.1. Penjelasan Dataset untuk WbFTL

Dataset	Simbol	Jumlah *Instances*	Jumlah *Features*	Jumlah Label	Keterangan
Office-Amazon	A	958	800	10 dari 31	Hasil pemrosesan SURF untuk gambar yang diambil dari *online merchant*, dengan berbagai *canonical viewpoint* dan resolusi medium yang diambil dari sebuah studio foto.
Office-DSLR	D	157	800	10 dari 31	Hasil pemrosesan SURF untuk gambar yang diambil menggunakan kamera DSLR pada lingkungan aslinya dengan pencahayaan alami, mempunyai resolusi tinggi (4288×2848)dan *noise* yang rendah.
Office-Webcam	W	295	800	10 dari 31	Hasil pemrosesan SURF untuk gambar yang diambil menggunakan *webcam* sederhana, beresolusi rendah (640×480), mengandung banyak *noise* dan *white balance artifacts*.
Caltech-256	C	1123	800	10 dari 256	Hasil pemrosesan SURF untuk gambar *online* dengan beragam resolusi, karena mempunyai beragam latar belakang, pencahayaan, ukuran, dan jenis kamera yang berbeda pada saat pengambilan gambar.

Gambar 4.1. Contoh Gambar Asli Dataset Amazon

4.2. Metode State-of The-Art

Seluruh pengujian unjuk kerja pada eksperimen ini dilakukan dengan menghitung akurasi, seperti yang juga dilakukan pada penelitian *feature-based transfer learning* lainnya. Pengujian akurasi ini merupakan tahapan untuk melihat posisi unjuk kerja WbFTL dibandingkan dengan metode feature-based transfer learning sebelumnya. Berdasarkan hasil pengujian juga akan ditentukan beberapa ide perbaikan dan pengembangan yang dapat dilakukan. Pengujian akurasi dilakukan per tahapan learning, serta dengan membandingkan dengan *state-of-the-art* dari penelitian *feature-based* transfer learning. Beberapa *state-of-the-art* dari *feature-based* transfer learning yang digunakan, yaitu: TCA, TJM, JDA, BDA, MEDA, VDA, SA, Coral.

4.3. Algoritma WbFTL

Detail dari algoritma WbFTL dapat dilihat pada Tabel 4.2 berikut.

Tabel 4.2. Algoritma WbFTL

Algorithm WbFTL
01: **procedure** transformasi *feature* antar domain
02: **for** $i \leftarrow 1$ to n do
03: $F_{Si} \leftarrow mean(f_{Si})$
04: $F_{Ti} \leftarrow mean(f_{Ti})$
05: $f'_i \leftarrow euclidean\ distance(F_{Si}, F_{Ti})$

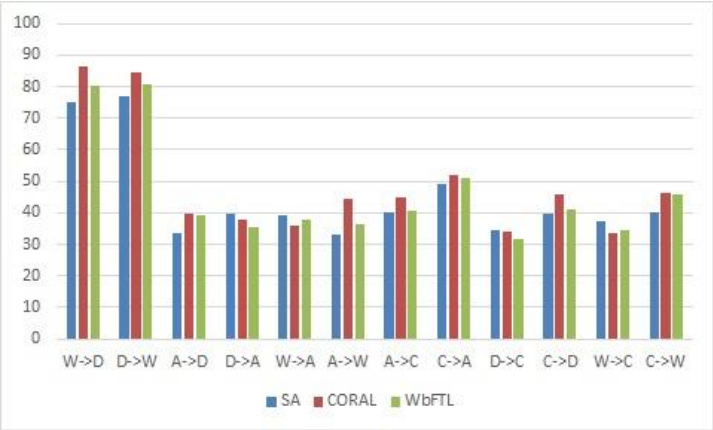
06:	**end for**
07:	return (f')
08: **end procedure**	
09: **procedure** transformasi *feature* dengan *threshold*	
10:	$Th \leftarrow mean(f')$
11:	$\dot{f}' \leftarrow f' < Th$
12:	**for** $i \leftarrow 1$ to k do
13:	$W(\dot{f}_i) \leftarrow \frac{\frac{1}{k}\sum_{i=1}^{k} \dot{f}_i}{\dot{f}_i}$
14:	$F_{STrans1(i)} \leftarrow W(\dot{f}_i) \times f_{Si}$
15:	$F_{TTrans1(i)} \leftarrow W(\dot{f}_i) \times f_{Ti}$
16:	**end for**
17:	return $(F_{STrans1}, F_{TTrans1})$
18: **end procedure**	
19: **procedure** transformasi *feature* dengan ANOVA	
20:	**for** $i \leftarrow 1$ to k do
21:	**for** $j \leftarrow 1$ to C do
22:	$F_{STrans2(i,j)} \leftarrow ANOVA\left(F_{STrans1(i)}\right)$
23:	**end for**
24:	**end for**
25:	return $(F_{STrans2})$
26: **end procedure**	
27: **procedure** inter-cluster class label	
28:	**for** $i \leftarrow 1$ to C do
29:	$h_{C(yi)} \leftarrow mean(F_{STrans2})$
30:	**end for**
31:	**for** $i \leftarrow 1$ to n_T do
32:	**for** $j \leftarrow 1$ to C do
33:	$D_{(i,j)} \leftarrow euclidean\ distance(F_{TTrans1(i)}, h_{C(yj)})$
34:	**end for**
35:	$dist(x_{Ti}, h_{C(yj)}) \leftarrow \min\ [D_{(i,j)}]$
36:	$h_{C(yi)} \leftarrow j$
37:	**end for**
38:	**for** $i \leftarrow 1$ to k do
39:	$F_{TTrans3} \leftarrow \left[\frac{\left(F_{TTrans1(i)} \times h_{C(yi)}\right)}{\left(dist\left(x_{Ti}, h_{C(yi)}\right)\right)^2}\right]$
40:	**end for**
41:	return $(F_{TTrans3})$
42: **end procedure**	

4.4. Hasil Eksperimen dengan Classifier SVM

Hasil perbandingan dapat dilihat di Gambar 4.2 dan Gambar 4.3, serta Tabel 4.3 dan Tabel 4.4. Gambar 4.2 merupakan visualisasi dari Tabel 4.3, yang menggambarkan perbandingan hasil akurasi WbFTL dengan metode *feature-transfer learning* sebelumnya yang termasuk pada kategori *easy*.

Tabel 4.3. Perbandingan Akurasi WbFTL dengan Easy Feature-based Transfer Learning

No	Dataset	Akurasi (%)		
		SA (2013)	CORAL (2016)	WbFTL (2023)
1	W->D	75,16(2)	**86,6(1)**	84,62(2)
2	D->W	**76,95(1)**	84,7(2)	**76,53(1)**
3	A->D	33,76(11)	39,5(9)	39,1(7)
4	D->A	39,87(6)	37,7(10)	34,38(10)
5	W->A	39,25(8)	36(11)	38,77(8)
6	A->W	33,22(12)	44,4(8)	37,41(9)
7	A->C	39,98(5)	45,1(7)	41,71(6)
8	C->A	49,27(3)	52,1(3)	50,57(3)
9	D->C	34,55(10)	33,8(11)	31,91(12)
10	C->D	39,49(7)	45,9(6)	44,23(5)
11	W->C	37,17(9)	33,7(12)	34,58(11)
12	C->W	40(4)	46,4(4)	45,58(4)
Rata-Rata		44,89	48,83	**46,62**
Peringkat		*3*	*1*	*2*

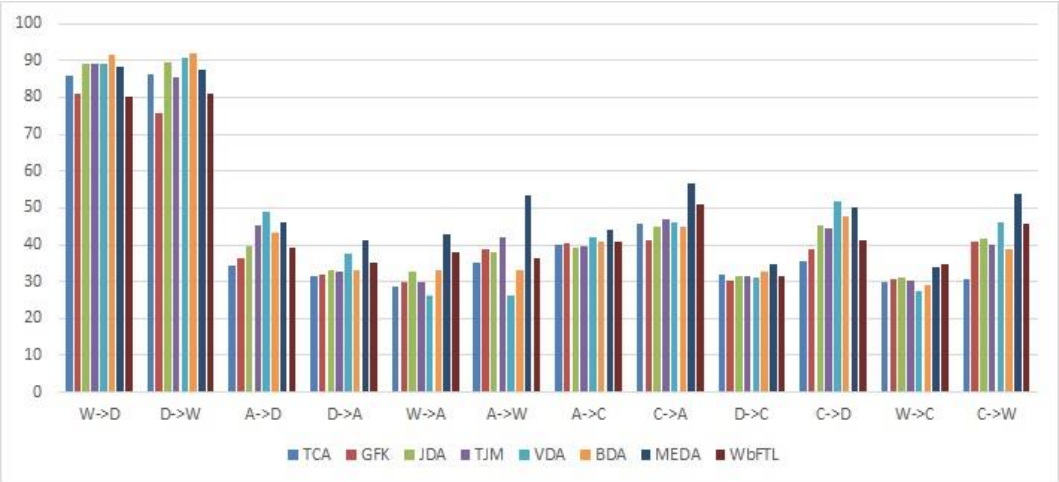


Gambar 4.2. Perbandingan Akurasi WbFTL dengan Easy Feature-based Transfer Learning per Pasangan Dataset

Gambar 4.3 merupakan visualisasi dari Tabel 4.4, yang menggambarkan perbandingan hasil akurasi WbFTL dengan metode *feature-transfer learning* sebelumnya yang termasuk pada kategori iteratif.

Tabel 4.4. Perbandingan Akurasi WbFTL dengan Iterative Feature-based Transfer Learning.

No	Dataset	Akurasi (%)							
		TCA (2011)	GFK (2012)	JDA (2013)	TJM (2014)	VDA (2017)	BDA (2018)	MEDA (2018)	WbFTL (2023)
1	**W->D**	85,99(2)	**80,89(1)**	89,17(2)	**89,2(1)**	89,18(2)	91,72(2)	**88,5(1)**	84,62(2)
2	**D->W**	**86,44(1)**	75,59(2)	**89,49(1)**	85,4(2)	**90,85(1)**	**91,86(1)**	87,5(2)	**76,53(1)**
3	**A->D**	34,39(7)	36,31(8)	39,49(6)	45,2(4)	48,81(4)	43,31(6)	45,9(7)	39,1(7)
4	**D->A**	31,42(9)	32,05(9)	33,09(9)	32,8(9)	37,68(8)	33,09	41,2(10)	34,38(10)
5	**W->A**	28,81(12)	29,75(12)	32,78(10)	30(12)	26,1(12)	32,99(10)	42,7(9)	38,77(8)
6	**A->W**	35,25(6)	38,98(6)	37,97(8)	42(6)	26,1(11)	32,99(9)	53,2(5)	37,41(9)
7	**A->C**	40,07(4)	40,25(5)	39,36(7)	39,5(8)	42,21(7)	40,78(7)	43,9(8)	41,71(6)
8	**C->A**	45,82(3)	41,02(3)	44,78(4)	46,8(3)	46,14(5)	44,89(5)	56,5(3)	50,57(3)
9	**D->C**	32,06(8)	30,28(11)	31,52(11)	31,4(10)	31,26(9)	32,5(11)	34,9(11)	31,91(12)
10	**C->D**	35,67(5)	38,85(7)	45,22(3)	44,6(5)	51,59(3)	47,77(3)	50,3(6)	44,23(5)
11	**W->C**	29,92(11)	30,72(10)	31,17(12)	30,2(11)	27,6(10)	28,94(12)	34(12)	34,58(11)
12	**C->W**	30,51(10)	40,68(4)	41,69(5)	40(7)	46,1(6)	38,64(8)	53,9(4)	45,58(4)
Rata-Rata		43,03	42,95	46,31	46,4	46,97	46,62	52,7	**46,62**
Peringkat		*8*	*7*	*5*	*4*	*2*	*3*	*1*	*3*



Gambar 4.3. Perbandingan Akurasi WbFTL dengan Iterative Feature-based Transfer Learning per Pasangan Dataset

4.5. Hasil Eksperimen dengan Classifier Logistic Regression

Sementara pada saat menggunakan classifier yang berbeda, yaitu logistic regression, hasil akurasi yang didapat sedikit berbeda. Hasil perbandingan akurasi dapat dilihat pada Tabel 4.5 dan Tabel 4.6. Tabel 4.5 merupakan perbandingan hasil akurasi WbFTL dengan metode *feature-transfer learning* sebelumnya yang termasuk pada kategori *easy*. Tabel 4.6 merupakan perbandingan perbandingan hasil akurasi WbFTL dengan metode feature-transfer learning sebelumnya yang termasuk pada kategori iteratif.

Tabel 4.5. Perbandingan Akurasi WbFTL dengan Logistic Regression pada Easy Feature-based Transfer Learning

No	Dataset	Akurasi (%)		
		SA (2013)	CORAL (2016)	WbFTL (*Logistic Regression*)
1	W->D	75,16(2)	**86,6(1)**	64.29(1)
2	D->W	**76,95(1)**	84,7(2)	49.53(9)
3	A->D	33,76(11)	39,5(9)	54.67(4)
4	D->A	39,87(6)	37,7(10)	50.38(8)
5	W->A	39,25(8)	36(11)	58.26(2)
6	A->W	33,22(12)	44,4(8)	52.64(6)
7	A->C	39,98(5)	45,1(7)	54.37(5)
8	C->A	49,27(3)	52,1(3)	42.78(12)
9	D->C	34,55(10)	33,8(11)	51.92(7)
10	C->D	39,49(7)	45,9(6)	43.13(10)
11	W->C	37,17(9)	33,7(12)	55.6(3)
12	C->W	40(4)	46,4(4)	42.85(11)
Rata-Rata		**44,89**	**48,83**	**51.70**
Peringkat		*3*	*2*	*1*

Hasil penelitian yang ditampilkan pada Tabel 4.5 dan Tabel 4.6 memperlihatkan bahwa akurasi WbFTL mengalami peningkatan pada saat digunakan bersama dengan classifier logistic regression, dibandingkan saat digunakan dengan SVM. Sehingga dapat mengungguli Coral, walaupun masih 1% dibawah MEDA. Hal ini menunjukkan bahwa WbFTL berjalan dengan baik, dan hasil yang lebih baik dapat diperoleh pada saat digunakan bersama classifier yang tepat.

Tabel 4.6. Perbandingan Akurasi WbFTL dengan Logistic Regression pada Iterative Feature-based Transfer Learning

No	Dataset	Akurasi (%)							
		TCA (2011)	GFK (2012)	JDA (2013)	TJM (2014)	VDA (2017)	BDA (2018)	MEDA (2018)	WbFTL (LR)
1	**W->D**	85,99(2)	**80,89(1)**	89,17(2)	**89,2(1)**	89,18(2)	91,72(2)	**88,5(1)**	64.29(1)
2	**D->W**	**86,44(1)**	75,59(2)	**89,49(1)**	85,4(2)	**90,85(1)**	**91,86(1)**	87,5(2)	49.53(9)
3	**A->D**	34,39(7)	36,31(8)	39,49(6)	45,2(4)	48,81(4)	43,31(6)	45,9(7)	**54.67(4)**
4	**D->A**	31,42(9)	32,05(9)	33,09(9)	32,8(9)	37,68(8)	33,09	41,2(10)	**50.38(8)**
5	**W->A**	28,81(12)	29,75(12)	32,78(10)	30(12)	26,1(12)	32,99(10)	42,7(9)	**58.26(2)**
6	**A->W**	35,25(6)	38,98(6)	37,97(8)	42(6)	26,1(11)	32,99(9)	53,2(5)	52.64(6)
7	**A->C**	40,07(4)	40,25(5)	39,36(7)	39,5(8)	42,21(7)	40,78(7)	43,9(8)	**54.37(5)**
8	**C->A**	45,82(3)	41,02(3)	44,78(4)	46,8(3)	46,14(5)	44,89(5)	56,5(3)	42.78(12)
9	**D->C**	32,06(8)	30,28(11)	31,52(11)	31,4(10)	31,26(9)	32,5(11)	34,9(11)	**51.92(7)**
10	**C->D**	35,67(5)	38,85(7)	45,22(3)	44,6(5)	51,59(3)	47,77(3)	50,3(6)	43.13(10)
11	**W->C**	29,92(11)	30,72(10)	31,17(12)	30,2(11)	27,6(10)	28,94(12)	34(12)	**55.6(3)**
12	**C->W**	30,51(10)	40,68(4)	41,69(5)	40(7)	46,1(6)	38,64(8)	53,9(4)	42.85(11)
Rata-Rata		**43,03**	**42,95**	**46,31**	**46,4**	**46,97**	**46,62**	**52,7**	**51.70**
Peringkat		*7*	*8*	*6*	*5*	*3*	*4*	*1*	*2*

4.6. Perbandingan Hyperparameter

Efisiensi dari penggunaan parameter pada WbFTL ini dapat terjadi karena setiap penambahan parameter akan memerlukan upaya untuk proses *tunning*. *Tunning hyperparameter* ini dilakukan dalam rangka mendapatkan hasil akurasi terbaik. Proses *tunning hyperparameter* ini tidak perlu dilakukan pada WbFTL yang tidak menggunakan parameter apapun. Oleh karena itu WbFTL ini dikatakan lebih efisien daripada metode *state-of-the-art*. Perbandingan hyperparameter yang digunakan dapat dilihat pada Tabel 4.7

Tabel 4.7. Perbandingan Jumlah Parameter WbFTL dengan Metode Feature-based Transfer Learning Sebelumnya

Metode	**TCA (2011)**	**JDA (2013)**	**TJM (2014)**	**VDA (2017)**	**BDA (2018)**	**MEDA (2018)**	**SA (2013)**	**CoRaL (2016)**	**WbFTL (2023)**
Akurasi (%)	43.03	46.31	46.4	46.97	46.62	52.7	45.9	48.83	46.6
Hyper parameter	μ	λ	λ	λ	λ, μ	λ, η, ρ	β_1, β_2	Tidak ada	Tidak ada

BAB 5
PENUTUP

5.1. Kesimpulan

Berdasarkan hasil eksperimen yang dilakukan, dapat diambil beberapa kesimpulan sebagai berikut:

A. Akurasi:

Hasil WbFTL memberikan akurasi 46,6% dengan menggunakan SVM sebagai *classifier*. Nilai akurasi ini baru dapat mengimbagi akurasi dari metode *feature-based transfer learning* yang menjadi *state-of-the-art*. Namun pada saat dilakukan pengembangan eksperimen dengan mengganti *classifier logistic regression*, nilai akurasi yang dihasilkan meningkat menjadi 51.7%. Nilai ini melebihi CoRaL, salah satu metode pada kelompok *easy feature-based transfer learning* yang mempunyai akurasi tertinggi pada kelompoknya. Sementara saat dibandingkan dengan kelompok *iterative feature-based transfer learning*, nilai yang dihasilkan ini tidak jauh berbeda dengan MEDA, yang mempunyai nilai akurasi rata-rata tertinggi pada kelompoknya.

Hasil yang diperoleh dalam penelitian ini juga menjawab seluruh pertanyaan penelitian. Yang mana dapat dikatakan bahwa akurasi klasifikasi yang baik dapat diperoleh dengan menerapkan pendekatan pemilihan dan pembobotan feature serta featura mapping. Bahkan akurasi klasifikasi yang baik pada cross-domain juga memungkinkan untuk dilakukan dengan cara yang sederhana, tanpa harus menggunakan banyak parameter dalam proses pemelajarannya.

B. Domain Asal (D_S):

Dengan melihat hasil akurasi dari masing-masing pasangan *cross-domain*, maka dapat dilihat bahwa hasil eksperimen memperlihatkan pada saat D_S berukuran besar, maka terdapat kecenderungan nilai akurasi juga baik. Nilai ini dapat dilihat pada saat dataset Caltech (C) yang menjadi D_S. Nilai akurasi juga cenderung tinggi pada saat jumlah *instance* pada D_S dan D_T lebih berimbang.

C. Pembuktian akan hipotesa penelitian:

Menggunakan hipotesa penelitian yang dituliskan, maka hasil yang didapat dari eksperimen WbFTL ini juga menjawab hipotesa penelitian yang dituliskan, bahwa:

- Representasi *feature* yang dibentuk dari kedekatan jarak antara *feature mean* pada *cross-domain* dapat dilakukan dan memberikan hasil yang baik, namun ini hanya untuk mengatasi perbedaan *marginal* pada *cross-domain*.
- Perbedaan *conditional distribution* pada *cross-domain* dapat dikurangi dengan memanfaatkan *unlabaled instances* pada D_T, melalui penggunaan rata-rata dan jarak, dapat dilihat dampaknya pada nilai akurasi akhir atau hasil akurasi pada transformasi *feature* kedua (F_{Trans2}).
- Peningkatan unjuk kerja metode terbukti dapat dilakukan melalui penerapan stategi pemilihan *feature* menggunakan rata-rata sebagai *threshold* dan varian untuk menilai *feature* yang informatif terhadap *class* label.

5.2. Kelebihan Metode WbFTL

Setelah melakukan analisa pada metode *feature-based* transfer learning yang menjadi *state-of-the-art* dan melakukan eksperimen untuk mengimplementasikan WbFTL, dapat disampaikan kelebihan dan kekurangan dari WbFTL. Kelebihan dari WbFTL adalah terletak pada kemudahan implementasi proses transformasi *feature* yang dilakukan, seperti:

- WbFTL menggunakan strategi pemilihan *feature* untuk memperkecil ruang pencarian sehingga dihasilkan metode yang lebih efisien dalam eksekusi.
- WbFTL menggunakan properti statistik yaitu rata-rata dan varian, serta jarak yang membuat metode ini lebih dinamis mengikuti kondisi data dan mempunyai tingkat skalabilitas yang baik.
- WbFTL juga memanfaatkan informasi pada *unlabaled instances* pada domain tujuan D_T, untuk memberikan informasi lebih banyak pada domain tujuan. Sehingga label hasil prediksi lebih dapat mencerminkan domain tujuan D_T.

5.3. Kekurangan Metode WbFTL

Sedangkan kekurangan dari WbFTL meliputi nilai akurasi yang dihasilkan, seperti:

- Akurasi yang dihasilkan oleh WbFTL dengan menggunakan SVM sebagai classifier baru dapat menyamai metode *feature-based transfer learning* sebelumnya, belum dapat mengungguli. Namun pada pengembangan eksperimen dengan menggunakan *logistic regression* sebagai classifier, hasil akurasi rata-rata WbFTL dapat mengungguli metode sebelumnya, kecuali untuk MEDA. Dibandingkan dengan MEDA, akurasi WbFTL masih 1% dibawah MEDA.
- Belum memperhatikan dampak dari masalah *imbalance instances* pada dataset yang digunakan.

5.4. Keterbatasan Metode WbFTL

Dalam pelaksanaan penelitian ini juga ditemukan beberapa keterbatasan, seperti:

- Penelitian WbFTL ini tidak menghitung dan membandingkan waktu (dalam satuan unit waktu) sebagai salah satu komponen yang dapat menjadi dasar evaluasi dan analisa.
- Penelitian WbFTL ini hanya menggunakan satu persamaan bobot dan belum dilakukan evaluasi dengan membandingkan hasilnya pada saat menggunakan persamaan bobot lainnya.
- Penelitian WbFTL ini juga hanya menggunakan satu persamaan jarak yaitu *Euclidean Distance* dalam proses transformasinya. Sehingga belum dilakukan evaluasi serta pengembangan penelitian dengan menggunakan persamaan jarak lainnya.

5.5. Saran Pengembangan

Melihat hasil analisa dan juga hasil eksperimen WbFTL yang telah dilakukan, maka ada beberapa saran untuk pengembangan selanjutnya (*future works*) yang dapat diterapkan, yaitu:

- Untuk meningkatkan akurasi metode, maka akan dilakukan dengan mencoba memodifikasi metode WbFTL terutama pada transformasi *feature* ketiga.

- Mengembangkan teknik pemilihan dan pembobotan *feature* yang digunakan pada metode WbFTL. Salah satu contohnya dengan mengikuti konsep *ordered weighted averaging* (OWA).
- Pengembangan penelitian juga dapat dilakukan dengan menerapkan *Vision Transformer* (ViT). ViT ini termasuk arsitektur *deep learning* dengan kebutuhan penggunaan sumber daya yang cukup kecil untuk proses training. Sehingga masih sejalan dengan konsep "sederhana" yang diusung oleh WbFTL.
- Implementasi WbFTL ini juga dapat dicoba diterapkan untuk tugas klasifikasi pada bidang lain seperti contohnya bidang agribisnis. Dalam bidang agribisnis, klasifikasi menggunakan *feature* dapat digunakan untuk mengetahui dengan cepat dan akurat kualitas produk hasil pertanian.

DAFTAR PUSTAKA

Borgwardt, K. M., Gretton, A., Rasch, M. J., Kriegel, H. P., Schölkopf, B., & Smola, A. J. (2006). Integrating structured biological data by Kernel Maximum Mean Discrepancy. *Bioinformatics*, *22*(14), 49–57. https://doi.org/10.1093/bioinformatics/btl242

Chattopadhyay, R., Sun, Q., Fan, W., Davidson, I., Panchanathan, S., & Ye, J. (2012). Multisource domain adaptation and its application to early detection of fatigue. *ACM Transactions on Knowledge Discovery from Data*, *6*(4). https://doi.org/10.1145/2382577.2382582

Cheplygina, V., Pena, I. P., Pedersen, J. H., Lynch, D. A., Sorensen, L., & De Bruijne, M. (2018). Transfer Learning for Multicenter Classification of Chronic Obstructive Pulmonary Disease. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1486–1496. https://doi.org/10.1109/JBHI.2017.2769800

Dai, W., & Xue, G. (n.d.). *Co-clustering based Classification for Out-of-domain Documents*. 210–219.

Davis, J., & Domingos, P. (2009). Deep transfer via second-order Markov logic. *Proceedings of the 26th International Conference On Machine Learning, ICML 2009*, 217–224.

Day, O., & Khoshgoftaar, T. M. (2017). A survey on heterogeneous transfer learning. *Journal of Big Data*, *4*(1). https://doi.org/10.1186/s40537-017-0089-0

Deng, W. Y., Lendasse, A., Ong, Y. S., Tsang, I. W. H., Chen, L., & Zheng, Q. H. (2019). Domain Adaption via Feature Selection on Explicit Feature Map. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(4), 1180–1190. https://doi.org/10.1109/TNNLS.2018.2863240

Fernando, B., Habrard, A., Sebban, M., & Tuytelaars, T. (2013). Unsupervised visual domain adaptation using subspace alignment. *Proceedings of the IEEE International Conference on Computer Vision*, *May 2014*, 2960–2967. https://doi.org/10.1109/ICCV.2013.368

Fernando, B., Habrard, A., Sebban, M., & Tuytelaars, T. (2014). *Subspace Alignment For Domain Adaptation*.

Frederick J Gravetter and Larry B. Wallnau. (2013). *Statistics for the Behavioral Sciences* (Vol. 9e). Wadworth Cengage Learning.

Gao, J., Fan, W., Jiang, J., & Han, J. (2008). Knowledge transfer via multiple model local structure mapping. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 283–291. https://doi.org/10.1145/1401890.1401928

Gong, B., Shi, Y., Sha, F., & Grauman, K. (2012). Geodesic flow kernel for unsupervised domain adaptation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2066–2073. https://doi.org/10.1109/CVPR.2012.6247911

Güneş, S., Polat, K., & Yosunkaya, Ş. (2010). Multi-class f-score feature selection approach to classification of obstructive sleep apnea syndrome. *Expert Systems with Applications*, *37*(2), 998–1004. https://doi.org/10.1016/j.eswa.2009.05.075

Han, H., Huang, M., Zhang, Y., Yang, X., & Feng, W. (2019). Multi-label learning with label specific features using correlation information. *IEEE Access*, *7*, 11474–11484. https://doi.org/10.1109/ACCESS.2019.2891611

Jiang, J., & Zhai, C. (2007). *Instance Weighting for Domain Adaptation in NLP Existing Work on Domain Adaptation • Existing work Outline*. 1–20.

Jiang, S., Xu, Y., Wang, T., Yang, H., Qiu, S., Yu, H., & Song, H. (2019). Multi-label metric transfer learning jointly considering instance space and label space distribution divergence. *IEEE Access*, *7*, 10362–10373. https://doi.org/10.1109/ACCESS.2018.2889572

John Blitzer Ryan McDonald Fernando Pereira. (2006). Domain Adaptation with Structural Correspondence Learning. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 120–128. https://doi.org/10.3115/1610075.1610094

Kouw, W. M., & Loog, M. (2018). *An introduction to domain adaptation and transfer learning*.

Kouw, W. M., & Loog, M. (2021). A Review of Domain Adaptation without Target Labels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(3), 766–785. https://doi.org/10.1109/TPAMI.2019.2945942

Kurniawan, F. (2019). *Deteksi Korban Selamat di Tengah Bencana Menggunakan Drone*. Sindonews.Com. https://jateng.sindonews.com/berita/10478/3/deteksi-korban-selamat-di-tengah-bencana-menggunakan-drone

Li, S., Chi Harold, L., Ding, Z., & Su, L. (2020). *Discriminative Transfer Feature and Label Consistency for Cross-Domain Image Classification*. IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Long, M., Wang, J., Ding, G., Pan, S. J., & Yu, P. S. (2014). Adaptation regularization: A general framework for transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *26*(5), 1076–1089. https://doi.org/10.1109/TKDE.2013.111

Long, M., Wang, J., Ding, G., Shen, D., & Yang, Q. (2012). *Transfer Learning with Graph Co-Regularization Graph Co-Regularized Transfer Learning*.

Long, M., Wang, J., Ding, G., Shen, D., & Yang, Q. (2014). Transfer learning with graph co-regularization. *IEEE Transactions on Knowledge and Data Engineering*, *26*(7), 1805–1818. https://doi.org/10.1109/TKDE.2013.97

Long, M., Wang, J., Ding, G., Sun, J., & Yu, P. S. (2013a). Transfer feature learning with joint distribution adaptation. *Proceedings of the IEEE International Conference on Computer Vision*, 2200–2207. https://doi.org/10.1109/ICCV.2013.274

Long, M., Wang, J., Ding, G., Sun, J., & Yu, P. S. (2013b). Transfer feature learning with joint distribution adaptation. *Proceedings of the IEEE International Conference on Computer Vision*, 2200–2207. https://doi.org/10.1109/ICCV.2013.274

Long, M., Wang, J., Ding, G., Sun, J., & Yu, P. S. (2014). Transfer joint matching for unsupervised domain adaptation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1410–1417. https://doi.org/10.1109/CVPR.2014.183

Maity, R. (1977). Statistical methods in hydrology and Hydroclimatology. In *Statistical methods in hydrology*. https://doi.org/10.1201/9780429423116-36

Manikandan, S. (2011). Measures of central tendency: The mean. *Journal of Pharmacology and Pharmacotherapeutics*, *2*(2), 140–142. https://doi.org/10.4103/0976-500X.81920

Mihalkova, L., Huynh, T., & Mooney, R. J. (2007). Mapping and revising Markov logic networks for transfer learning. *Proceedings of the National Conference on Artificial Intelligence*, *1*(July), 608–614.

Pan, S. J., Kwok, J. T., & Yang, Q. (2008a). Transfer learning via dimensionality reduction. *Proceedings of the National Conference on Artificial Intelligence*, *2*, 677–682.

Pan, S. J., Kwok, J. T., & Yang, Q. (2008b). Transfer learning via dimensionality reduction. *Proceedings of the National Conference on Artificial Intelligence*, *2*, 677–682.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Pham, T. N., Tran, L. V., & Dao, S. V. T. (2020). Early Disease Classification of Mango Leaves Using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection. *IEEE Access*, *8*, 189960–189973. https://doi.org/10.1109/ACCESS.2020.3031914

Saad, E., Paprzycki, M., Ganzha, M., Badica, A., Badica, C., Fidanova, S., Lirkov, I., & Ivanovic, M. (2022). *Generalized Zero-Shot Learning for Image Classification—Comparing Performance of Popular Approaches*. 1–16. https://doi.org/10.3390/info13120561

Sanodiya, R. K., Mathew, A., Mathew, J., & Khushi, M. (2020). Statistical and Geometrical Alignment using Metric Learning in Domain Adaptation. *Proceedings of the International Joint Conference on Neural Networks*. https://doi.org/10.1109/IJCNN48605.2020.9206877

Sevani, N., Hermawan, I., & Jatmiko, W. (2019). Feature selection based on f-score for enhancing CTG data classification. *Proceedings: CYBERNETICSCOM 2019 - 2019 IEEE International Conference on Cybernetics and Computational Intelligence: Towards a Smart and Human-Centered Cyber World*, 18–22. https://doi.org/10.1109/CYBERNETICSCOM.2019.8875656

Shao, L., Zhu, F., & Li, X. (2015). Transfer learning for visual categorization: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(5), 1019–1034. https://doi.org/10.1109/TNNLS.2014.2330900

Sharma, M., & Batra, A. (2014). Analysis of Distance Measures in Content Based Image Retrieval. *Global Journal of Computer Science and Technology*, *14*(2), 11–16.

Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok, and Q. Y. (2011). *Domain Adaptation via Transfer Component Analysis*. *22*(2), 199–210. https://doi.org/10.1109/TNN.2010.2091281

Sun, B., Feng, J., & Saenko, K. (2016). Return of frustratingly easy domain adaptation. *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, *1*, 2058–2065.

Tahmoresnezhad, J., & Hashemi, S. (2017a). Visual domain adaptation via transfer feature learning. *Knowledge and Information Systems*, *50*(2), 585–605. https://doi.org/10.1007/s10115-016-0944-x

Tahmoresnezhad, J., & Hashemi, S. (2017b). Visual domain adaptation via transfer feature learning. *Knowledge and Information Systems*, *50*(2), 585–605. https://doi.org/10.1007/s10115-016-0944-x

Van Opbroek, A., Achterberg, H. C., Vernooij, M. W., & De Bruijne, M. (2019). Transfer learning for image segmentation by combining image weighting and kernel learning. *IEEE Transactions on Medical Imaging*, *38*(1), 213–224. https://doi.org/10.1109/TMI.2018.2859478

Wang, J., Chen, Y., Feng, W., Han, Y. U., Huang, M., & Yang, Q. (2020). Transfer learning with dynamic distribution adaptation. *ACM Transactions on Intelligent Systems and Technology*, *11*(1). https://doi.org/10.1145/3360309

Wang, J., Chen, Y., Hao, S., Feng, W., & Shen, Z. (2017). Balanced distribution adaptation for transfer learning. *Proceedings - IEEE International Conference on Data Mining, ICDM*, *2017-Novem*, 1129–1134. https://doi.org/10.1109/ICDM.2017.150

Wang, J., Chen, Y., Hu, L., Peng, X., & Yu, P. S. (2018). Stratified Transfer Learning for Cross-domain Activity Recognition. *IEEE International Conference on Pervasive Computing and Communications (PerCom)*.

Wang, J., Chen, Y., Yu, H., Huang, M., & Yang, Q. (2019). Easy transfer learning by exploiting intra-domain structures. *Proceedings - IEEE International Conference on Multimedia and Expo*, *2019-July*, 1210–1215. https://doi.org/10.1109/ICME.2019.00211

Wang, J., Feng, W., Chen, Y., Huang, M., Yu, H., & Yu, P. S. (2018a). Visual Domain Adaptation with Manifold Embedded. *ACM Internation Conference on Multimedia*, 402–410. https://doi.org/10.1145/3240508.3240512

Wang, J., Feng, W., Chen, Y., Huang, M., Yu, H., & Yu, P. S. (2018b). Visual domain adaptation with manifold embedded distribution alignment. *MM 2018 - Proceedings of the 2018 ACM Multimedia Conference*, *1*, 402–410. https://doi.org/10.1145/3240508.3240512

Weiss, K., Khoshgoftaar, T. M., & Wang, D. D. (2016a). A survey of transfer learning. In *Journal of Big Data* (Vol. 3, Issue 1). Springer International Publishing. https://doi.org/10.1186/s40537-016-0043-6

Weiss, K., Khoshgoftaar, T. M., & Wang, D. D. (2016b). A survey of transfer learning. In *Journal of Big Data* (Vol. 3, Issue 1). Springer International Publishing. https://doi.org/10.1186/s40537-016-0043-6

Xia, R., Zong, C., Hu, X., & Cambria, E. (2015). Feature ensemble plus sample selection: Domain adaptation for sentiment classification. *IJCAI International Joint Conference on Artificial Intelligence*, *2015-Janua*(Ijcai), 4229–4233.

Xiaomeng, L., Yu, L., Fu, C.-W., Fang, M., & Heng, P.-A. (2020). Revisiting metric learning for few-shot image classification. *Elsevier*. https://doi.org/10.1016/j.neucom.2020.04.040

Xu, Y., Pan, S. J., Xiong, H., Wu, Q., Luo, R., Min, H., & Song, H. (2017). A Unified Framework for Metric Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, *29*(6), 1158–1171. https://doi.org/10.1109/TKDE.2017.2669193

Zhang, J., Li, W., & Ogunbona, P. (2017). Joint geometrical and statistical alignment for visual domain adaptation. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*(May), 5150–5158. https://doi.org/10.1109/CVPR.2017.547

Zhang, Y., Hu, X., & Fang, Y. (2010). Logistic Regression for Transductive Transfer Learning from Multiple Sources. In L. Cao, J. Zhong, & Y. Feng (Eds.), *Advanced Data Mining and Applications* (Vol. 6441, pp. 175–182). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-17313-4_18

Zhong, X., Guo, S., Shan, H., Gao, L., Xue, D., & Zhao, N. (2018). Feature-Based Transfer Learning Based on Distribution Similarity. *IEEE Access*, *6*, 35551–35557. https://doi.org/10.1109/ACCESS.2018.2843773

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2019). *A Comprehensive Survey on Transfer Learning*. *10*, 1–27.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, *109*(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555