



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL



Nama : Gery Jonathan Manurung

NIM : 211402137

Judul diajukan oleh\* : ☐ Dosen  
☒ Mahasiswa

Bidang Ilmu (tulis dua bidang) : 1. Data Science and Intelligent System  
2. Computer Graphics and Vision

Uji Kelayakan Judul\*\* : ☒ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Umayra Ramadhani Putri Nasution S.TI., M.Kom.  
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Paraf Calon Dosen Pembimbing I

Calon Dosen Pembimbing II: Dr. Sawaluddin M.IT

Medan, 4 Februari 2025

Ka. Laboratorium Penelitian,

\* Centang salah satu atau keduanya

\*\* Pilih salah satu

(Fanindia Purnamasari S.TI., M.IT)

NIP. 198908172019032023



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

RINGKASAN JUDUL YANG DIAJUKAN

\*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	<b>Deteksi Deepfake pada Wajah dalam Video dengan Cross-Attention Multi-Scale Vision Transformer dan EfficientNet</b>
Latar Belakang dan Penelitian Terdahulu	<p><i>Deepfake</i> adalah istilah yang berasal dari gabungan dua kata, yaitu "<i>deep learning</i>" dan "<i>fake</i>". Teknologi ini mengacu pada video atau konten digital yang telah dimanipulasi secara hiper-realistis untuk menampilkan perbuatan seseorang yang sebenarnya tidak pernah terjadi (Westerlund, 2019). <i>Deepfake</i> bekerja dengan menggunakan jaringan saraf buatan, khususnya <i>Generative Adversarial Networks</i> (GANs), untuk menganalisis kumpulan besar data seperti ekspresi wajah, gerak tubuh, suara, hingga infleksi suara seseorang. Teknologi ini menggabungkan kemampuan pemetaan wajah dengan algoritma AI untuk menukar wajah seseorang dalam video dengan wajah orang lain.</p> <p>Kemajuan teknologi <i>deepfake</i> yang semakin mudah diakses dan canggih telah menimbulkan ancaman besar, khususnya di Indonesia. <i>Deepfake</i> kini digunakan dalam berbagai modus penipuan digital, seperti <i>social engineering</i>, <i>account takeover</i>, pencurian identitas, dan pemalsuan dokumen. Menurut riset VIDA (2024), "<i>Where's The Fraud - Protecting Indonesia Business from AI Generated Fraud</i>", kasus penipuan berbasis <i>deepfake</i> di Indonesia melonjak hingga 1550% antara tahun 2022 dan 2023, menunjukkan betapa seriusnya dampak teknologi ini terhadap ekosistem digital.</p> <p>Berbagai pendekatan digunakan untuk mendeteksi <i>deepfake</i>, yang dapat dikategorikan menjadi metode berbasis <i>deep learning</i>, <i>machine learning</i>, dan model statistik. Metode berbasis <i>deep learning</i> mendominasi karena kemampuannya mengekstraksi fitur langsung dari data, dengan model seperti CNN (XceptionNet, ResNet, EfficientNet) dan RNN (LSTM) menunjukkan akurasi rata-rata 89,73% dan AUC sebesar 0,917 pada dataset seperti FaceForensics++ dan Celeb-DF. Sebaliknya, metode <i>machine learning</i>, seperti SVM dan <i>Random Forest</i>, lebih sederhana tetapi kurang efektif dalam menangani dataset kompleks. Model statistik, seperti <i>Kullback-Leibler divergence</i>, meskipun memiliki kegunaan tertentu, jarang digunakan karena keterbatasan dalam menangani manipulasi kompleks. Namun, meski unggul, metode <i>deep learning</i> menghadapi tantangan seperti kebutuhan data besar dan penurunan performa pada <i>dataset</i> di luar distribusi latihannya (Rana et al., 2022).</p> <p>Salah satu tantangan utama dalam deteksi <i>deepfake</i> adalah ketidakmampuan model untuk melakukan generalisasi dengan baik pada data eksternal, sering kali disebabkan oleh <i>overfitting</i> akibat data yang kurang beragam. Das et al. (2021) mengidentifikasi bahwa <i>oversampling dataset</i> menjadi salah satu penyebab utama dan memperkenalkan metode <i>Face-Cutout</i>, yaitu teknik augmentasi data yang memotong sebagian kecil area wajah pada data latih. Teknik ini meningkatkan variasi data dan membantu jaringan saraf konvolusional (CNN) menghindari <i>overfitting</i>. Dengan metode ini, performa deteksi <i>deepfake</i> meningkat hingga 15,2%–35,3% pada berbagai arsitektur CNN, tanpa modifikasi signifikan pada <i>pipeline</i> deteksi. Selain itu, metode ini terbukti fleksibel karena tidak bergantung pada jenis <i>dataset</i> tertentu.</p> <p>Durasi video yang panjang juga menjadi tantangan, karena dapat menyebabkan komputasi berlebihan selama pelatihan model. Bonettini et al. (2021) menemukan bahwa penggunaan 32 frame per video memberikan keseimbangan antara efisiensi komputasi dan performa model,</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

mengurangi risiko *overfitting* tanpa meningkatkan *validation loss* secara signifikan dibandingkan 15 frame. Untuk efisiensi, BlazeFace *extractor* digunakan untuk mengekstraksi wajah dari setiap frame dengan ukuran input  $224 \times 224$  piksel. Namun, penelitian ini tidak memanfaatkan *Long Short-Term Memory* (LSTM) untuk menganalisis hubungan temporal antar *frame*, sehingga potensi mendeteksi inkonsistensi gerakan dan ekspresi wajah belum sepenuhnya dieksplorasi.

Penelitian yang dilakukan oleh Haq (2021) berfokus pada pengembangan model deteksi video *deepfake* menggunakan *dataset* Celeb-DF(V2). Dengan mengekstraksi 2000 *frame* wajah dari video asli dan *deepfake*, peneliti menggunakan model XceptionNet yang dilatih dengan teknik pemrosesan gambar seperti *Gaussian Filter* dan *Local Binary Pattern* (LBP). Hasilnya menunjukkan performa yang signifikan, dengan nilai *Area Under Curve* (AUC) sebesar 0.87 dan akurasi 0.79, mengungguli model ResNet-50 yang juga diuji dalam penelitian ini. Temuan ini menegaskan potensi besar XceptionNet dalam mendeteksi manipulasi wajah pada video *deepfake*. Namun, penelitian ini masih dapat dikembangkan lebih lanjut untuk meningkatkan performa dan generalisasi model, terutama dalam menghadapi perkembangan teknologi *deepfake* yang semakin kompleks dan canggih.

*Self-attention mechanism* pada *Vision Transformer* (ViT) telah membawa terobosan besar dalam deteksi *deepfake*. Mekanisme ini memungkinkan model mempelajari hubungan global antar *pixel*, sehingga dapat menangkap informasi kontekstual tingkat tinggi untuk mendeteksi manipulasi visual secara lebih akurat. Penelitian oleh Khormali dan Yuan (2022) menunjukkan bahwa *framework* DFDT berbasis ViT mencapai akurasi deteksi hingga 99,41% pada *FaceForensics++*, 99,31% pada *Celeb-DF (V2)*, dan 81,35% pada *WildDeepfake*. Kemampuan multi-skala ViT memungkinkan deteksi manipulasi pada berbagai skala dengan efisiensi tinggi. Namun, pendekatan ini menghadapi tantangan seperti kebutuhan *dataset* besar untuk menghindari *overfitting*, risiko *attention collapse*, serta minimnya eksplorasi ViT dibandingkan metode berbasis CNN yang lebih mapan.

Kombinasi arsitektur *EfficientNet* dan *Vision Transformer* (ViT) telah menunjukkan hasil yang superior dalam analisis citra. Penelitian oleh Duong et al. (2021) membuktikan bahwa konfigurasi ViT\_Base\_Eff\_B1\_224 mampu mencapai akurasi 97,72% dalam deteksi tuberkulosis dari citra X-ray dada. Keunggulan ini didukung oleh efisiensi fitur ekstraksi dari *EfficientNet* dan mekanisme *self-attention* pada ViT yang memungkinkan pemodelan kompleksitas hubungan spasial dalam data citra. Strategi transfer learning menggunakan model pra-latih seperti *AdvProp* dan *Noisy Student* juga mempercepat proses pembelajaran dan meningkatkan generalisasi model tanpa memerlukan *dataset* yang sangat besar.

Pendekatan CrossViT menawarkan keunggulan tambahan dibandingkan ViT standar. Dengan arsitektur *dual-branch* dan mekanisme *cross-attention*, CrossViT mampu menangkap fitur multi-skala secara efektif, memperkaya representasi fitur, dan meningkatkan akurasi klasifikasi hingga 82,8% pada *ImageNet1K* (Chen et al., 2021). Selain efisiensi komputasi, CrossViT juga unggul dalam fleksibilitas *transfer learning*. Namun, kompleksitas arsitektur CrossViT dapat menjadi tantangan dalam proses pelatihan, terutama dalam skenario dengan *dataset* besar dan kompleks.

Berdasarkan penelitian yang telah dilakukan deteksi *deepfake* menghadapi tantangan seperti keterbatasan generalisasi pada data eksternal, kebutuhan *dataset* besar, risiko *overfitting*, dan tingginya beban komputasi. CNN efektif dalam ekstraksi fitur lokal namun terbatas dalam menangkap hubungan spasial kompleks, sementara ViT unggul dalam pemodelan hubungan global tetapi rentan terhadap *attention collapse* dan membutuhkan *dataset* besar.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Mengatasi kendala tersebut, kombinasi *Cross-Attention Multi-Scale Vision Transformer* (Cross-ViT) dan EfficientNet diusulkan. Cross-ViT menangkap artefak manipulasi pada berbagai skala spasial dengan mekanisme *cross-attention*, sementara EfficientNet memberikan efisiensi komputasi melalui scaling parameter yang optimal. Dengan teknik augmentasi seperti *Face Cutout*, kombinasi ini menjaga akurasi yang baik dan meningkatkan generalisasi serta efisiensi model dalam mendeteksi *deepfake*.

Penelitian Terdahulu

No.	Penulis	Judul	Tahun	Keterangan
1.	Sowmen Das, Selim Seferbekov, Arup Datta, Md. Saiful Islam, Md. Ruhul Amin	Towards Solving the DeepFake Problem : An Analysis on Improving DeepFake Detection using Dynamic Face Augmentation	2021	Penelitian ini mengusulkan metode augmentasi data baru, <i>Face-Cutout</i> , yang menggunakan informasi landmark wajah untuk memotong area tertentu pada gambar, guna mengurangi <i>overfitting</i> dan meningkatkan generalisasi model. Hasil menunjukkan peningkatan performa deteksi <i>DeepFake</i> dengan penurunan <i>LogLoss</i> sebesar 15.2% hingga 35.3% dibandingkan teknik oklusi lainnya.
2.	Nicolò Bonettini, Edoardo Daniele Cannas, Sara Mandelli, Luca Bondi, Paolo Bestagini, Stefano Tubaro	Video Face Manipulation Detection Through Ensemble of CNNs	2021	Penelitian ini menggunakan pendekatan ensemble CNN yang dilengkapi <i>attention mechanism</i> dan strategi pelatihan siamese untuk mendeteksi manipulasi wajah dalam video. Hasil menunjukkan performa deteksi yang unggul dibandingkan <i>baseline</i> pada <i>dataset</i> FF++ dan DFDC, dengan efisiensi analisis 4.000 video dalam waktu kurang dari 9 jam.
3.	Javid Al Haq	Klasifikasi Cepat Model Xceptionnet Dan Resnet-50 Pada Video Deepfake Menggunakan Local Binary Pattern	2021	Penelitian ini mengembangkan metode deteksi <i>deepfake</i> dengan memanfaatkan <i>Local Binary Pattern</i> (LBP) dan model XceptionNet serta ResNet-50. Hasil menunjukkan bahwa XceptionNet memiliki performa lebih baik dengan AUC 0.87 dan akurasi validasi 79%, unggul dalam klasifikasi video asli maupun <i>deepfake</i> dibandingkan ResNet-50.





**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI**

**UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI**

**PROGRAM STUDI S1 TEKNOLOGI INFORMASI**

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: [tek.informasi@usu.ac.id](mailto:tek.informasi@usu.ac.id) | Laman: <http://it.usu.ac.id>

	4.	Aminollah Khormali, Jiann-Shiun Yuan	DFDT: An End-to-End DeepFake Detection Framework Using Vision Transformer	2022	Penelitian ini mengembangkan DFDT, framework deteksi deepfake berbasis <i>Vision Transformer</i> yang memodelkan hubungan global antar piksel dan memanfaatkan mekanisme re-attention. DFDT mencapai akurasi tinggi hingga 99.41% pada FaceForensics++ dan menunjukkan generalisasi kuat antar <i>dataset</i> , membuktikan efektivitasnya dalam mendeteksi berbagai jenis manipulasi gambar.
	5.	Linh T. Duong, Nhi H. Le, Toan B. Tran, Vuong M. Ngo, Phuong T. Nguyen	Detection of tuberculosis from chest X-ray images: Boosting the performance with vision transformer and transfer learning	2021	Penelitian ini memanfaatkan kombinasi EfficientNet dan <i>Vision Transformer</i> (ViT) dengan <i>transfer learning</i> untuk mendeteksi tuberkulosis dari citra rontgen dada. Model yang dikembangkan mencapai akurasi 97.72% dan nilai AUC tinggi, menunjukkan performa unggul dalam membedakan kasus positif dan negatif dari model lain serta potensi aplikasi luas dalam analisis citra medis.
	6.	Chun-Fu Richard, Chen, Quanfu Fan, Rameswar Panda	CrossViT: Cross-Attention Multi-Scale Vision Transformer for Image Classification	2021	Penelitian ini memperkenalkan CrossViT, model <i>Vision Transformer</i> yang menggunakan dua cabang dengan ukuran <i>patch</i> berbeda dan <i>cross-attention</i> untuk memproses informasi multi-skala. CrossViT-18+T2T mencapai akurasi top-1 83.0% pada <i>ImageNet1K</i> .
<b>Rumusan Masalah</b>		Teknologi <i>deepfake</i> terus berkembang pesat, menghasilkan manipulasi visual yang semakin realistis dan mudah diakses oleh masyarakat luas. Hal ini menimbulkan berbagai permasalahan, terutama dalam penyebaran informasi palsu, penipuan digital, dan ancaman terhadap privasi serta keamanan data. Banyak kasus penyalahgunaan <i>deepfake</i> yang digunakan untuk kepentingan kriminal, seperti pencurian identitas, pemalsuan dokumen, hingga manipulasi opini publik. Di sisi lain, upaya untuk mendeteksi <i>deepfake</i> masih menghadapi berbagai tantangan, mulai dari akurasi yang belum optimal karena <i>overfitting</i> hingga efisiensi sistem yang belum memadai dalam memproses data dalam jumlah besar. Oleh karena itu, diperlukan sistem deteksi <i>deepfake</i> berbasis <i>website</i> yang dapat bekerja secara efisien dan juga akurat untuk mengidentifikasi konten manipulatif serta mengurangi dampak negatifnya di berbagai aspek kehidupan.			



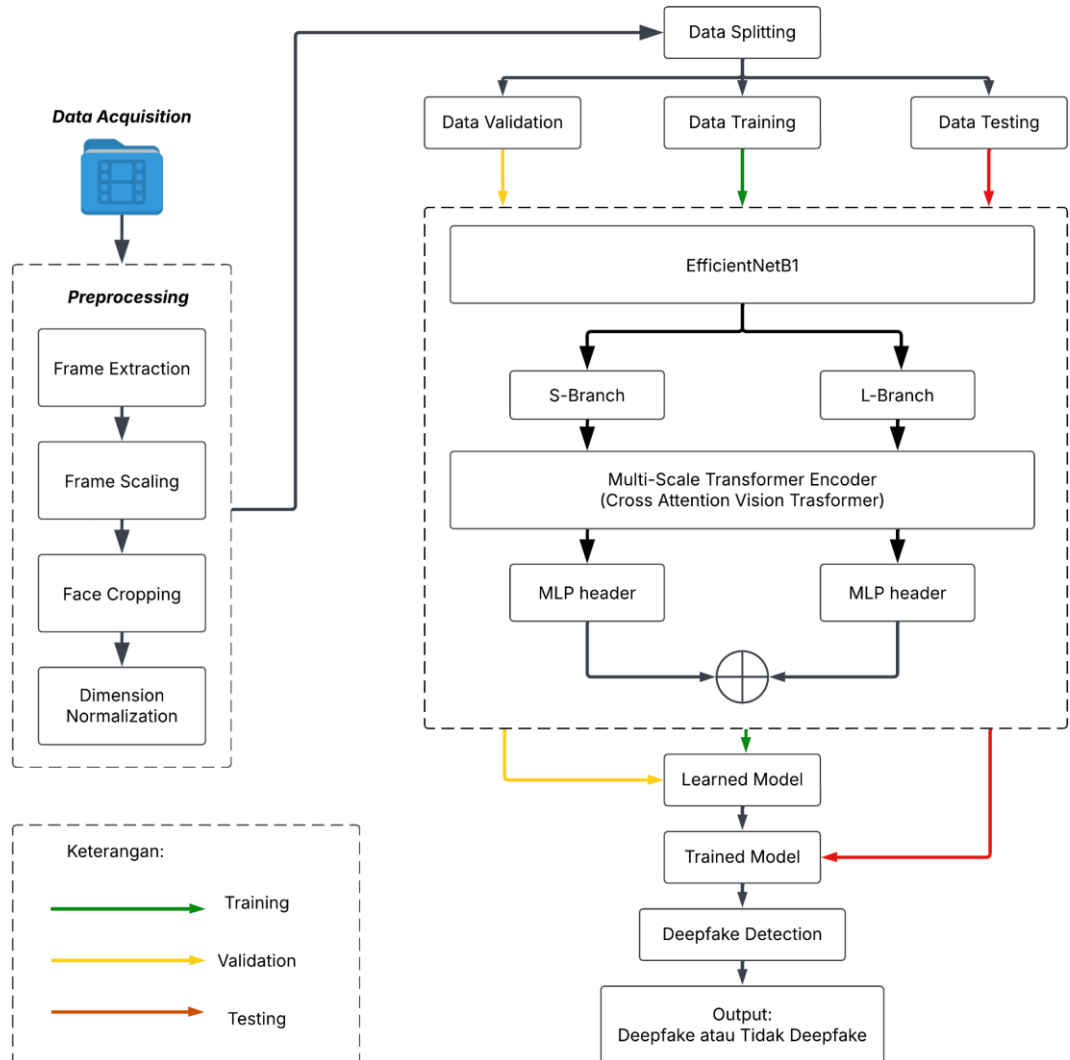
KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Metodologi



Tahapan Penelitian:

1. **Data Acquisition**

Data yang digunakan dalam penelitian ini berasal dari *dataset* Celeb-DF v2 yang memiliki label *ground truth* untuk video *deepfake* dan tidak *deepfake*. *Dataset* ini dipilih karena kualitas manipulasi yang realistis, menjadikannya tantangan yang sesuai untuk evaluasi model deteksi *deepfake*.

2. **Preprocessing**

Tahap *preprocessing* dilakukan untuk mempersiapkan data mentah agar sesuai dengan kebutuhan model. Langkah-langkah *preprocessing* mencakup:

- Frame Extraction:** Mengambil sebanyak 32 *frame* representatif dari setiap video untuk mewakili kontennya.
- Frame Scaling:** Dilakukan penyesuaian ukuran berdasarkan resolusi:
  - 2x *Scaling* untuk video dengan sisi terlebar kurang dari 300 piksel.
  - Tidak ada *rescaling* untuk video dengan sisi terlebar antara 300 hingga 1000 piksel.
  - 0.5x *Scaling* untuk video dengan sisi terlebar lebih dari 1000 piksel.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	<ul style="list-style-type: none"><li>● 0.33x <i>Scaling</i> untuk video dengan sisi terlebar lebih dari 1900 piksel.</li><li>c. <i>Face Cropping</i>: Menggunakan MTCNN (<i>Multi-task Cascaded Convolutional Networks</i>) untuk mendeteksi dan memotong wajah dari setiap <i>frame</i>.</li><li>d. <i>Dimension Normalization</i>: Setiap <i>frame</i> disesuaikan ukurannya agar kompatibel dengan dimensi input EfficientNetB1, yaitu 240x240 piksel.</li></ul> <p>3. <b>Data Splitting</b> <i>Dataset</i> dibagi menjadi tiga bagian dengan proporsi 50% untuk training data, 20% untuk validation, dan 30% untuk testing. Pembagian ini memastikan evaluasi yang seimbang untuk model.</p> <p>4. <b>Modeling</b> Tahapan ini mencakup desain dan integrasi model yang digunakan untuk deteksi <i>deepfake</i>. EfficientNetB1 berfungsi sebagai <i>feature extractor</i> yang menghasilkan dua cabang pemrosesan:</p> <ul style="list-style-type: none"><li>a. <i>S-Branch (Small Patch Processing)</i>: Menggunakan <i>patch</i> berukuran kecil untuk menangkap informasi lokal seperti artefak kecil pada wajah akibat manipulasi <i>deepfake</i>.</li><li>b. <i>L-Branch (Large Patch Processing)</i>: Menggunakan <i>patch</i> berukuran besar untuk menangkap informasi global dari <i>frame</i> yang dapat membantu memahami konteks manipulasi secara keseluruhan.</li></ul> <p>Sebagai bagian dari augmentasi data, dilakukan teknik <i>face cutout</i> pada training data. Teknik ini memotong sebagian wajah secara acak untuk mencegah <i>overfitting</i> dan meningkatkan kemampuan generalisasi model.</p> <p><i>Output</i> dari kedua cabang (<i>S-Branch</i> dan <i>L-Branch</i>) diintegrasikan menggunakan mekanisme <i>cross-attention</i> pada CrossViT (<i>Cross-Attention Vision Transformer</i>). Mekanisme ini memungkinkan interaksi langsung antara informasi lokal dan global, yang meningkatkan kemampuan model dalam mendeteksi manipulasi <i>deepfake</i>. Setiap cabang menghasilkan <i>logits</i> terpisah yang kemudian dijumlahkan untuk menghasilkan probabilitas akhir dari klasifikasi. Model ini memprediksi apakah sebuah video adalah <i>deepfake</i> atau tidak <i>deepfake</i>.</p> <p>5. <b>Output</b> Hasil dari proses <i>training</i> adalah website dengan model deteksi <i>deepfake</i> yang mampu menerima input video dan memberikan prediksi apakah video tersebut merupakan <i>deepfake</i> atau bukan. Model ini dapat digunakan dalam aplikasi deteksi <i>deepfake</i> secara otomatis.</p>
Referensi	<p>Bonettini, N., Bondi, L., Cannas, E. D., Bestagini, P., Mandelli, S., &amp; Tubaro, S. (2020). Video face manipulation detection through ensemble of CNNs. <i>Proceedings - International Conference on Pattern Recognition</i>. <a href="https://doi.org/10.1109/ICPR48806.2021.9412711">https://doi.org/10.1109/ICPR48806.2021.9412711</a></p> <p>Chen, C. F., Fan, Q., &amp; Panda, R. (2021). CrossViT: Cross-Attention Multi-Scale Vision Transformer for Image Classification. <i>Proceedings of the IEEE International Conference on Computer Vision</i>. <a href="https://doi.org/10.1109/ICCV48922.2021.00041">https://doi.org/10.1109/ICCV48922.2021.00041</a></p> <p>Das, S., Seferbekov, S., Datta, A., Islam, M. S., &amp; Amin, M. R. (2021). Towards Solving the DeepFake Problem : AAn Analysis on Improving DeepFake Detection using Dynamic</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

**PROGRAM STUDI S1 TEKNOLOGI INFORMASI**

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

- Face Augmentation. *Proceedings of the IEEE International Conference on Computer Vision, 2021-October*. <https://doi.org/10.1109/ICCVW54120.2021.00421>
- Duong, L. T., Le, N. H., Tran, T. B., Ngo, V. M., & Nguyen, P. T. (2021). Detection of tuberculosis from chest X-ray images: Boosting the performance with vision transformer and transfer learning. *Expert Systems with Applications, 184*.  
<https://doi.org/10.1016/j.eswa.2021.115519>
- Haq, J. Al. (2021). Klasifikasi cepat model xceptionnet dan ResNet-50 pada video deepfake menggunakan local binary pattern. In *Repository.Uinjkt.Ac.Id*.  
<https://repository.uinjkt.ac.id/dspace/handle/123456789/58537>
- Khormali, A., & Yuan, J. S. (2022). DFDT: An End-to-End DeepFake Detection Framework Using Vision Transformer. *Applied Sciences (Switzerland), 12*(6).  
<https://doi.org/10.3390/app12062953>
- Rana, M. S., Nobil, M. N., Murali, B., & Sung, A. H. (2022). Deepfake Detection: A Systematic Literature Review. In *IEEE Access* (Vol. 10).  
<https://doi.org/10.1109/ACCESS.2022.3154404>
- VIDA. (2024). Penipuan deepfake Indonesia melonjak 1550%, begini cara VIDA memeranginya. VIDA.  
<https://vida.id/id/pressrelease/penipuan-deepfake-indonesia-melonjak-1550-begini-cara-vida-memeranginya>
- Westerlund, M. (2019). The emergence of deepfake technology: A review. In *Technology Innovation Management Review* (Vol. 9, Issue 11).  
<https://doi.org/10.22215/TIMREVIEW/1282>

Medan, 4 Februari 2025  
Mahasiswa yang mengajukan,

(Gery Jonathan Manurung)

NIM. 211402137