BAB I
PENDAHULUAN

1.1. Latar Belakang

Peningkatan kualitas citra (*Image Enhancement*) merupakan salah satu topik yang cukup menjadi perhatian dalam bidang *Computer Vision* hingga saat ini yang dibuktkan dengan meningkatnya jumlah publikasi artikel di berbagai jurnal internasional. Setelah dilakukan riset dengan *keyword* pencarian *Low-light Image Enhancement* di jurnal *open-access* terkemuka yaitu *Arxiv* dan MDPI mencatat bahwa tahun 2015 hingga tahun 2025 terjadi peningkatan eksponensial yang sangat signifikan. Hal tersebut tidak terlepas dari peningkatan kebutuhan akan pengambilan gambar digital yang berkualitas tinggi di berbagai kondisi. Salah satu kondisi yang menjadi tantangan dalam menghasilkan gambar dengan kualitas tinggi adalah kondisi tingkat pencahayaan yang rendah (*Low-Light*), dikarenakan kompeksnya untuk meningkatkan visibilitas dan menjaga kealamian gambar secara visual [1]. Tantangan ini diakui pada ajang internasional *NTIRE Challenge*, di mana tim-tim teratas mencapai PSNR di atas 25 dB dalam tugas *Low-Light Image Enhancement*, menandakan bahwa topik ini memang layak untuk menjadi perhatian dan menunjukkan kemajuan signifikan dalam pengembangan teknologi yang inovatif [2]. Sedangkan peningkatan kualitas citra digital dengan tingkat pencahayaan rendah banyak diaplikasikan diberbagai bidang, diantaranya pada aplikasi pengawasan keamanan, fotografi malam, dan pengintaian militer [3].

Sementara itu, di Indonesia berdasarkan data dari PUSIKNAS POLRI yang dikutip oleh akurat.co mengungkapkan bahwa kasus pencurian berat (curat) merupakan kasus yang paling banyak terjadi dengan total 44.671 kasus sepanjang tahun 2025 [4]. Sedangkan menurut Direktur Reserse Kriminal umum Polda Metro Jaya, Kombes Pol Hengky Haryadi tindak kejahatan sering terjadi di jam rawan antara pukul 00.00 hingga 04.00 shubuh [5]. Hal ini merujuk pada pemanfaatan kamera pengawsan yang intensif untuk mengungkap bukti digital pada kondisi minim cahaya, disinilah *Low-light Image Enhancement* berperan. Mengingt juga *device* atau kamera yang ada sekarang ini belum mampu menghasilkan citra digital

berkualitas seperti yang diharapkan. Yang dimana saat pengambilan gambar pada pencahayaan rendah sering mengandung blur (buram), *noise*, kurang kontras, dan tidak mempu menangkap detail citra, sehingga akan membatasi fungsionalitas dan estetika dari citra tersebut.

Selain dari sisi kemampuan *device* atau kamera dalam menangkap citra nyata menjadi digital, pengembangan teknologi tambahan yang`lebih canggih berdasarkan teknik pemrosesan gambar tradisional juga telah berkontribusi untuk menghasilkan gambar digital yang berkualitas, seperti *Histogram Equalization* (HE), *Contrast- Limited Adaptive Histogram Equalization* (CLAHE), *Gamma Ajustment* (GA), dan metode tradisional lainya [6]-[8]. Namun, metode tradisional seperti *Histogram Equalization* sering kali menyebabkan peningkatan pencahayaan berlebihan dan hilangnya detail, sehingga menghasilkan tampilan yang tidak jelas dan tidak alami dalam peningkatan gambar di cahaya rendah [9]. Hal tersebut disebabkan karena teknologi tradisional tersebut kurang fleksibel dengan banyaknya karakteristik citra digital yang dihasilkan, sehingga memang bukan sebuah pilihan yang bijak ketika menggunakan teknik tradisional ini pada banyak gambar yang memiliki variasi karakteristik berbeda beda, seperti jenis *noise*, jenis blur, intensitas cahaya yang juga variatif. Selain metode tradisional diatas, ada metode tradisional yang lebih adaptif terhadap kondisi pencahayaan gambar yaitu *Imadjust* yang diintegrasikan pada *software MatLab.* Dimana *Imadjust* menggunakan *scaling min-max*, *non-linear gamma transformation*, dan *linear transformation* dengan *hyperparameter* yang bisa diatur manual sesuai dengan kondisi gambar. Namun metode ini tentunya akan tetap kesulitan ketika berhadapan dengan otomatisasi proses secara *real time* untuk gambar dengan kondisi yang sangat variatif seperti kamera pengawasan atau *autonomous driving.* Karena mayoritas metode tradisional masih harus mengatur *hyperparameter* manual dan belum bisa menekan *noise* serta kualitas gambar keseluruhan, maka metode tradisional ini belum cukup menyelesaikan masalah *Low-light Image Enhancement* pada sistem *real time* atau tugas yang sangat mempertimbangkan kualitas gambar, Untuk itu dibutuhkan sebuah metode yang bisa mengoptimalkan kedua aspek tersebut yaitu kualitas gambar dan ketahanan terhadap variasi data.

Sementara pada beberapa tahun terakhir ini, terjadi pengembangan metode yang cukup masif yang didominasi oleh pendekatan berbasis *Deep Learning,* di mana banyak strategi pembelajaran, struktur jaringan, fungsi kerugian, data pelatihan, dll. yang mendukung fleksibilitas dan skalabilitas model yang lebih luas [10]. Bebebrapa algoritma *deep learning* yang sering digunakan oleh beberapa pengembang dan peneliti diantaranya adalah *Autoencoder*, *Convolutional Neural Networks* (CNN), hingga pengembangan dengan model *generative* seperti *Generative Adversarial Networks* (GANs) yang dipercaya cukup mampu menangani permasalahan peningkatan kualitas citra terutama pada kondisi minim cahaya. Tidak hanya sampai pada titik itu, saat ini sudah cukup banyak algoritma *deep learning* yang dimodifikasi untuk tugas spesifik, salah satunya adalah *U-Net* yang pertama kali diperkenalkan pada tahun 2015 sebagai model untuk segmentasi citra biomedis [11]. U-Net adalah arsitektur *deep learning* yang menggunakan komponen seperti *encoder*, *decoder*, *skip connection*, *bridge network*, *loss function*, dan kombinasinya untuk segmentasi gambar vaskular dan non-vaskular [12]. Namun U-Net merupakan arsitektur yang sangat fleksibel dan mudah untuk dimodifikasi, disesuaikan dengan berbagai tugas *computer vision* lainnya, salah satunya adalah *low-light image enhancement*.

Hingga saat ini para peneliti sangat antusias untuk mengembangkan dan memodifiksi jaringan U-Net ini salah satunya, pada tahun 2019 tim peneliti dari China melakukan modifikasi arsitektur U-Net untuk *Low-light Image Enhancement* [13]. Penelitiannya berangkat dari problem *downsampling* biasa yang tidak cukup baik dalam menjaga detail gambar sehingga sulit dipelajari, oleh karena itu mereka memperkenalkan Transformasi *Wavelet*. Transformasi *Wavelet* memungkinkan gambar didekomposisi menjadi 4 sub-gambar yaitu *sub-band* (LL), detail horizontal (LH), detail vertikal (HL), dan detail diagonal (HH) kemudian resolusi gambar diturunkan menjadi setengah dari gambar aslinya. sehingga gambar memiliki informasi yang lebih kaya dari berbagai sudut frekuensi gambar walaupun resolusi diturunkan setengahnya. Mekanisme *downsampling* dan *upsampling* seperti ini dinamakan dengan DWT (*Discrete Wavelet Transform*) dan IDWT (*Inverse Discrete Wavelet Transform*). Model ini dibekali dengan fungsi loss L1 (*Mean*

Absolute Error) dan *Perceptual loss* dengan model *pre-trained*. Model ini dilatih dengan *Adam optimizer* selama 3500 *epoch* dengan penjadwalan *learning rate* secara bertahap, dan hasil akhirnya model ini mampu mencapai PSNR = 28.92 dan SSIM = 0.851 pada benchmark SID dataset [14]. Secara kualitas gambar, Pendekatan ini memang menghasilkan gambar yang halus, noise pada gambar terreduksi dengan sangat baik. Namun memiliki kelemahan dalam ketajaman detail lokal, dan menghasilkan grid semu (garis - garis) pada gambar yang dihasilkan.

Di sisi lain, tetap pada tahun 2019 juga ada penelitian yang mengembangkan Arsitektur U-Net dengan mengusulkan RRCU *Module* (*Recurrent Residual Convolutional Units*) dan *Multi-ways dilated bottleneck* [15]. Arsitektur U-Net dimodifikasi dengan mengganti *double convolution* pada U-Net Standard dengan RRCU *module* untuk melakukan ekstraksi fitur lebih baik. Kemudian *downsample* hanya dilakukan dua kali, dan dua sisanya diganti dengan *dilated convolution*. *Baottleneck layer* diterapkan operasi *Multi-ways dilated convolution* dan average pooling. kemudian model dilatih dengan fungsi loss *Mean Absolute Error* (MAE) dan *Adam Optimizer* selama 5000 *epoch*. Gambar yang dihasilkan oleh model ini cukup baik dengan skor PSNR = 30.29 dan SSIM = 0.7776 pada pengujian *benchmark* SID dataset [14]. Secara visual model ini sangat unggul dalam mempertahankan akurasi warna namun secara detail struktural gambar masih mengasilkan artefak dan kurang tajam. Hipotesis dari penyebab masalah ini adalah terlalu banyak mengekstrak informasi global dengan *dilated convolution* dan tidak ada fungsi *loss* khusus untuk menangani detail lokal dengan lebih baik.

Kemudian pada tahun 2024 masih dari peneliti China menyadari bahwa model U-Net ini cukup mahal secara memori dikarenakan mekanisme *skip connection* disetiap levelnya [16]. Kemudian dari masalah tersebut mereka mengusulkan U-Net-- yang mengunggulkan efisiensi memori pada *skip connection*. Model U-Net-- ini dibekali *Multiscale Information Agregation Module* (MSIAM) untuk mereduksi memori pada *skip connection* dan tetap mempertahankan pemetaan fitur di berbagai level [16]. *Information Enhancement Module* (IEM) juga diusulkan untuk memperkaya fitur yang sudah di agregasi. pada penelitian ini telah dicoba diberbagai tugas pemrosesan gambar seperti *Image deblurring* dengan

perolehan skor PSNR = 40.01 dan SSIM = 0.962. terlepas dari keunggulan efisiensi memori, model ini memiliki kekurangan dalam akurasi warna Ketika diterapkan pada tugas *low-light image enhancement*, yang memiliki indikasi disebabkan karena proses reduksi channel menjadi channel tunggal pada modul MSIAM.

Dengan berbagai inovasi yang diusulkan oleh para peneliti sebelumnya diatas, penulis memberikan usulan model baru bernama LL-DSV-UNet sebagai penyempurnaan dari model - model sebelumnya agar mendukung keberlanjutan penelitian. Penulis mengintegrasikan beberapa komponen pada LL-DSV-UNet, diantaranya *Residual Dense Block Network* (RDB-Net) sebagai alternatif *double convolution* dari U-Net Standard dan RRCU dari modifikasi U-Net di paper [15]. RDB-Net dengan *multiscale feature extraction* yang disusun secara sekuensial dan koneksi yang kuat disetiap skala, maka akan menghasilkan *feature maps* yang lebih kaya dengan komputasi yang lebih ringan. Kemudian diterapkan *multilevel input* dengan mengkombinasikan antara *Averagepooling* dan *Maxpooling* sehingga proses ekstraksi fitur disetiap level lebih halus tanpa mereduksi struktur lokal secara signifikan. Karena *low-light image enhancement* merupakan tugas yang cukup sensitif terhadap variasi pencahayaan, maka diusulkan *Contextual Attention Module* (CAM) sebagai upaya peningkatan pencahayaan yang lebih adaptif. *Deep Supervision mechanism* yang terinspirasi model U-Net++ [17] juga diterapkan untuk memandu proses rekonstruksi dengan baik, namun secara implementasi penulis modifikasi dengan berbasis *multi-level fusion output* bernama MFDS yang nantinya hanya akan ada satu perhitungan *loss* terhadap hasil penggabungan *output multi-level* sehingga aliran gradien lebih stabil. Dan terakhir customisasi fungsi *loss* dengan mengukur diferensiasi setiap piksel menggunakan *Mean Absolute Error* (MAE), diferensiasi distribusi setiap *channel* warna dengan mengukur nilai *Mean Square Error* (MSE) dan *Standard Deviation Squared Error* (SDSE) untuk mendukung akurasi warna dari gambar dengan lebih baik, dan pengukuran diferensiasi kontras dan structural dari gambar dengan nilai rata- rata dan kovarian antara gambar target (*Ground Truth*) dan gambar prediksi.

Dengan menerapkan inovasi setiap komponen ini, LL-DSV-UNet dapat menghasilkan model *deep learning* baru yang lebih baik, dari sisi kualitas gambar

khususnya ketajaman strktural dan *naturalness*, serta mudah diadaptasikan untuk mendukung tugas peningkatan kualitas citra pada gambar dengan tingkat pencahayaan rendah pada berbagai domain spesifik. Selain daripada pengusulan arstektur model LL-DSV-UNet yang merupakan hasil dari modifikasi arsitektur U-Net [11], penulis juga akan merancang sebuah *Graphical User Interface* (GUI) untuk nantinya *user* dapat menggunakan model yang sudah dikembangkan ketika berkaitan dengan *Low-light Image Enhancement*. Untuk penggunaan yang efisien, sistem GUI ini dibangun menggunakan *Flask* sebagai *backend* Python dan HTML/CSS murni sebagai *frontend,* dengan komunikasi melalui HTTP POST *form*. Fitur yang utama dari system GUI ini adalah dukungan *multi-model* dengan bisa memilih opsi model inferensi pada menu *drop-down* dan user dapat mengunggah gambar input (*low-light*) dengan dukungan berbagai format gambar seperti JPG dan PNG. Desain GUI ini dibuat sangat sederhana, dikarenakan pada dasarnya *Low-light Image Enhancement* akan jauh lebih berguna Ketika diintegrasikan pada sistem lain sebagai langkah optimasi, seperti kasus segmentasi gambar pada citra yang gelap.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diungkapkan pada poin 1.1, maka rumusan masalah yang akan diselesaikan dalam penelitian ini disajikan seperti berikut :

1. Bagaimana cara merancang arsitektur *deep learning* berbasis U-Net yang fokus pada kualitas perseptual dan struktural gambar untuk meningkatkan kualitas gambar dengan tingkat pencahayaan rendah?
2. Bagaimna performa model *deep learning* yang diusulkan dalam meningkatkan kualitas gambar dengan tingkat pencahayaan rendah dibandingkan dengan model berbasis *deep learning* lainnya khususnya pada aspek perseptual dan struktural gambar?
3. Bagaimana Merancang *Graphical User Interface* (GUI) untuk proses inferensi peningkatan kualitas gambar dengan tingkat pencahayaan rendah menggunakan model yang diusulkan?

1.3 Batasan Masalah

Pembatasan suatu masalah digunakan untuk menghindari adanya penyimpangan maupun pelebaran pokok masalah agar penelitian tersebut lebih terarah dan memudahkan dalam pembahasan sehingga tujuan penelitian akan tercapai. Beberapa batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya fokus pada modifikasi arsitektur U-Net, modifikasi fungsi loss, dan evaluasi performa model yang diusulkan terhadap model *U- Net Standard* [11], *Wavelet U-Net* [13], *U-Net Modified* [15], dan *U-Net* -- [16].
2. Penelitian ini tidak menyediakan solusi *end to end* untuk permasalahan optimasi kamera pengawasan, namun hanya menyediakan komponen model *deep learning* saja.
3. Evaluasi performa model ditinjau dari aspek stabilitas, konvergensi, dan kualitas gambar yang dihasilkan - masing masing model berdasarkan *pixel wise intensity*, kesamaan struktural, dan kesamaan perseptual.

1.4 Tujuan Penelitian

Dengan didasarkan pada rumusan masalah yang diperinci seperti diatas, maka penelitian ini memiliki tujuan sebagai berikut :

1. Merancang dan mengembangkan arsitektur model *deep learning* berbasis U-Net yang fokus pada kualitas perseptual dan struktural gambar untuk meningkatkan kualitas gambar dengan tingkat pencahayaan rendah
2. Mengevaluasi performa dan efektivitas model *deep learning* berbasis U-Net yang diusulkan dalam meningkatkan kualitas gambar dengan tingkat pencahayaan rendah terhadap model *deep learning* lainnya khususnya pada aspek perseptual dan structural gambar
3. Mendesain dan merancang *Graphical User Interface* (GUI) untuk proses inferensi peningkatan kualitas gambar dengan tingkat pencahayaan rendah menggunakan model yang diusulkan.

1.5 Manfaat Penelitian

Hasil penelitian ini, diharapkan dapat memberikan mafaat diantaranya sebagai berikut :

1. Bagi bidang keilmuan

   Menemukan dan Mengetahui inovasi baru dalam perancangan model *deep learning* untuk *low-light image enhancement* berbasis arsitektur U-Net, serta bagaimana menerapkan teknologi *Deep Supervision* untuk mencapai performa model yang jauh lebih baik dan stabil.

2. Bagi Masyarakat

   Membantu mengurangi kasus kejahatan yang terjadi di tengah masyarakat melalui penguatan Lembaga Penegak Hukum dalam mengungkap bukti, sehingga akan tercipta kondisi yang aman dan tertib.

3. Bagi Pemerintah

   Membantu mempermudah POLRI atau Lembaga Penegak Hukum dalam mengungkap bukti digital atas kasus kejahatan yang terjadi di malam hari melalui kamera pengawasan.