

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Pemanfaatan teknologi yang terkomputerisasi oleh manusia selalu meninggalkan jejak yang tersimpan dalam bentuk data digital. Rekam jejak ini merupakan bukti perilaku dan karakteristik manusia sehingga dijadikan sebagai acuan pengembangan teknologi dan ilmu pengetahuan pada masa mendatang. Data digital yang umumnya dimanfaatkan oleh manusia meliputi teks, citra audio, citra visual, dan citra audio visual yang disimpan ke dalam suatu media penyimpanan. Banyaknya jumlah data yang tersedia dan diprediksi akan semakin bertambah membuat gaya hidup manusia semakin bergantung pada teknologi digital.

Jejak digital bersifat laten yang berarti data digital memiliki makna khusus dan hanya dapat diolah dengan prosedur khusus. Citra visual terbentuk dari penangkapan pantulan gelombang elektromagnetik benda yang berada depan lensa kamera monokuler dan kemudian disimpan dalam bentuk digital. Citra gambar monokuler memiliki informasi laten dengan semantik tertentu sehingga dapat dijadikan *input* yang bagus. Hasil akhir citra visual yang tercipta hanya memiliki informasi numerik mengenai warna, sedangkan informasi posisi benda saat penangkapan sudah hilang. Estimasi informasi laten dapat dilakukan komputer dengan membuat fungsi pemetaan dari suatu gambar terhadap suatu kondisi dengan memanfaatkan data dalam jumlah besar sebagai acuan.

Permodelan pembelajaran dalam atau *deep learning* dapat memetakan suatu domain ke domain lainnya secara mandiri menggunakan pembelajaran jaringan saraf tiruan dalam atau *deep neural network*. Pembelajaran dalam dapat dilakukan dengan komputasi mandiri yang sangat bergantung pada kuantitas dan kualitas data yang baik. Pembelajaran dalam menggunakan jaringan saraf tiruan dapat digunakan untuk mengembangkan teknologi khususnya di bidang visi komputer seperti melakukan estimasi pose tiga dimensi tubuh manusia yang terdapat dalam suatu gambar monokuler.

## 1.2 Rumusan Masalah

Permasalahan yang ingin diselesaikan adalah mengestimasi pose tubuh manusia yang bersifat laten dengan membuat sebuah fungsi pemetaan dari gambar yang ditangkap kamera monokuler ke koordinat setiap titik kunci pose menggunakan model *deep neural network*. Model yang telah terlatih mampu melakukan proses inferensi terhadap gambar baru sehingga dapat dijadikan sebagai acuan aplikasi yang mampu membaca pose tubuh manusia dari gambar monokuler.

## 1.3 Batasan Masalah

Penelitian ini menganggap setiap pose dua dimensi maupun pose tiga dimensi berada dalam koordinat lokal. Pose yang dihasilkan tidak bersifat *grounded* yang berarti pose tidak memperhitungkan lantai sebagai titik tengah. Setiap pose ditransformasi ke dalam sistem koordinat kamera dengan titik kunci pinggang sebagai titik koordinat tengah. Hal ini dilakukan karena pemetaan hanya menggunakan gambar dua dimensi tanpa informasi kedalaman titik kunci sehingga dapat mengesampingkan masalah kedalaman yang ambigu.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk membuat sebuah aplikasi yang dapat mengestimasi titik kunci dari sebuah citra visual datar dan melakukan transformasi ke pose lokal dua dimensi sehingga dapat dipetakan oleh jaringan saraf tiruan yang dimodelkan ke bentuk pose lokal tiga dimensi. Pose hasil juga divisualisasikan secara interaktif sehingga dapat digunakan untuk kepentingan yang sesuai.

## 1.5 Metode Penelitian

Penelitian dibagi menjadi tiga tahap besar terurut yang terdiri dari *data preprocessing*, pelatihan model jaringan saraf tiruan, dan visualisasi. Tahap pertama dan tahap ketiga tidak melibatkan pemelajaran mesin. Permasalahan

utama dari penelitian ini berada pada tahap kedua tentang pelatihan jaringan saraf tiruan untuk melakukan pemetaan pose dua dimensi ke pose tiga dimensi.

*Dataset* yang digunakan dalam penelitian ini adalah *Human3.6M* yang berisi 3,6 juta pose unik yang dilakukan oleh sebelas peraga profesional dan direkam menggunakan empat sudut kamera yang berbeda beserta dengan koordinat setiap titik kunci dari hasil penangkapan alat *motion capture* [11].

Pose dua dimensi dan tiga dimensi merupakan variabel yang relevan untuk masalah. Pada tahap *data preprocessing* akan dilakukan ekstraksi variabel ini menjadi bentuk numerik sehingga mudah untuk digunakan saat melakukan pelatihan jaringan saraf tiruan. Tahap selanjutnya akan dilakukan pembuatan, permodelan, pelatihan, dan evaluasi jaringan saraf tiruan dalam untuk pemetaan titik kunci yang kemudian dilanjutkan dengan percobaan model dengan video. Tahap terakhir menampilkan visualisasi gambar, pose dua dimensi, dan pose tiga dimensi.

Perangkat keras yang digunakan dalam penelitian ini adalah satu unit laptop dengan spesifikasi:

- CPU Intel Core I7 7700HQ
- Memori 24 GB DDR4
- GPU NVIDIA GTX 1060 6GB
- SSD NVME SAMSUNG 120 GB
- HDD SATA 1 TB

Perangkat lunak yang digunakan dalam penelitian ini adalah satu unit laptop dengan spesifikasi:

- Python 3.7
- Jupyter Lab
- Git

- GitHub
- LaTeX

## **1.6 Sistematika Penulisan**

Adapun sistematika penulisan yang digunakan dalam penulisan ini adalah sebagai berikut:

PENDAHULUAN, mengemukakan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

TINJAUAN PUSTAKA, menjelaskan kumpulan teori yang digunakan dalam mendukung proses penyelesaian program.

PENDEKATAN, mengemukakan langkah-langkah yang dicapai untuk membuat program.

HASIL DAN ANALISIS, mendalami hasil yang tercapai dengan analisis secara mendalam.

PENUTUP, mengulas lebih lanjut mengenai kesimpulan yang dapat ditarik dari hasil disertai dengan saran yang dapat menyempurnakan penelitian selanjutnya.