

Pengenalan Bentuk Tangan dengan Convolutional Neural Network (CNN)

Muhammad Zein Eryad¹, Kurniawan Nur Ramadhani², Anditya Arifianto³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹zersyad@students.telkomuniversity.ac.id, ²kurniawannr@telkomuniversity.ac.id

³anditya@telkomuniversity.ac.id

Abstrak — Pada saat ini terdapat beberapa sistem pengenalan gestur tangan yang memiliki tingkat komputasi dan upaya untuk persiapan perangkat yang cukup merepotkan. Tujuan utama penulis adalah membangun sistem yang mendekati pengenalan gestur / bentuk tangan yang memiliki tingkat komputasi dan upaya yang lebih rendah, serta mendapatkan hasil performa sistem pada saat proses pengujian. Sistem yang dibangun menggunakan *Convolutional Neural Network (CNN)* dengan *input* berupa citra dan *output* berupa label dari *input* tersebut. Pengembangan sistem ini sudah dilakukan hingga tahap pengujian, *model* yang dihasilkan mendapatkan akurasi klasifikasi sebesar 88% yang diuji dengan 2142 citra dan digambarkan dengan *confusion matrix* sebagai alat ukur performansi.

Kata kunci : bentuk tangan, cnn, confusion matrix.

1 Pendahuluan

1.1 Latar Belakang

Pengenalan gestur merupakan topik pada bidang ilmu *computer science and language technology* dengan tujuan mengenali arti dari bahasa tubuh manusia dengan menggunakan algoritma matematika. Gestur berasal dari *motion* atau *state* tubuh, gestur yang paling sering digunakan berasal dari bagian tangan atau wajah. Pada pengenalan gestur tangan, terdapat beberapa cara yang dapat dilakukan, yang pertama dapat menggunakan sinyal radar yang memancarkan aliran elektromagnetik [1], yang kedua dapat menggunakan kamera dan algoritma *computer vision* untuk menerjemahkan *motion* atau *state* tangan. Pengenalan gestur dapat dilakukan dengan tiga implementasi, yaitu dengan menggunakan perangkat sarung tangan, *3-dimensional location hand points* dan dengan *raw visual data*. Implementasi pertama yang mengharuskan menggunakan sarung tangan dengan kabel yang tersambung pada sarung tangan dan komputer meskipun memiliki kecepatan proses dan tingkat akurasi yang lebih baik. Implementasi kedua menggunakan *hand point extraction* yang mana membutuhkan biaya komputasi yang tinggi. Implementasi terakhir menggunakan sensor gambar, seperti kamera, sensor inframerah atau sensor kedalaman yang mana tergantung kepada pengguna [2].

Convolutional Neural Network (CNN) atau biasa disebut *ConvNet*, *CNN* mengekstrak fitur dari *input* yang berupa gambar lalu mengubah dimensi gambar tersebut menjadi lebih kecil tanpa merubah karakteristik gambar tersebut. *CNN* terdiri dari *neurons* yang memiliki bobot dan bias. Setiap *neuron* menerima inputan dan diteruskan dengan melakukan perkalian titik pada setiap *neuron* tersebut. Pada *layer* terakhir *CNN* masih memiliki *loss function* seperti *SVM/Softmax*. Dengan menggunakan sensor gambar seperti kamera, upaya yang digunakan untuk mempersiapkan perangkat jauh lebih mudah, misalnya seperti hampir setiap orang saat ini sudah memiliki perangkat kamera pada *smartphone*. Penelitian ini memberikan suatu inovasi berdasarkan adanya masalah pada pengenalan gestur dalam bidang *computing* yaitu pengenalan gestur tangan memiliki beberapa opsi, beberapa opsi tersebut memiliki *cost* dan *resource* yang tinggi, dan juga upaya mempersiapkan perangkat yang cukup merepotkan.

Dilihat dari permasalahan tersebut, penelitian ini mengembangkan suatu solusi yaitu dengan mengembangkan sistem pengenalan bentuk tangan dengan menggunakan *model* yang sudah dilatih sebelumnya, *model* dapat digunakan dalam perangkat pintar seperti *smartphone*. Penelitian ini menggunakan suatu metode yang dapat merealisasikan tujuan dari penelitian ini yaitu terdapat beberapa implementasi seperti implementasi dengan sarung tangan untuk mendapatkan bentuk tangan, implementasi kedua menggunakan *hand point extraction* dengan biaya komputasi yang tinggi, implementasi ketiga menggunakan sensor kamera. Penelitian ini menggunakan pengenalan bentuk tangan yang merupakan salah satu proses dari pengenalan gestur tangan. Pengembangan sistem untuk pengenalan bentuk tangan dapat dibantu dengan menggunakan salah satu metode yaitu *Convolutional Neural Network*. *CNN* adalah salah satu dari beberapa metode *Deep Learning* yang memiliki hasil paling signifikan dalam pengenalan citra [3]. *CNN* mempunyai kelebihan dibandingkan dengan metode lainnya yaitu dapat memproses

komputasi untuk melatih *model* sebelum melakukan pengujian, dengan itu saat pengujian dilakukan tidak memerlukan pelatihan yang berulang, *model* bisa digunakan dimana saja. Namun metode CNN memiliki kelemahan yaitu semakin banyak dataset yang digunakan pada proses pelatihan, semakin lama juga proses pelatihan [3] dan *model* yang dilatih berdasarkan dataset yang diberikan, oleh karena itu hasil dari pengenalan bentuk bergantung pada *input* (dataset) yang diberikan pada saat pelatihan. Penelitian ini mengharapkan rencana pembangunan sistem bentuk tangan dengan CNN dapat berhasil sesuai tujuan yang telah ditentukan.

1.2 Topik dan Batasannya

Topik yang diutamakan ialah klasifikasi gambar dengan menggunakan *Convolutional Neural Network*. Pada klasifikasi memiliki batasan untuk pengenalan tujuh bentuk tangan dengan masing-masing bentuk tangan diberikan label *Palm Oke, Fist, C, L, Index point* dan *Up*, pemilihan tujuh kelas diambil karena keterbatasan *resource* yang digunakan pada saat proses pelatihan model *CNN*.

1.3 Tujuan

Tujuan yang dicapai dalam pengerjaan tugas akhir ini adalah membangun sistem untuk mengenali bentuk tangan dan mendapatkan hasil dari kinerja performa dari sistem setelah dilakukan pengujian oleh data uji.

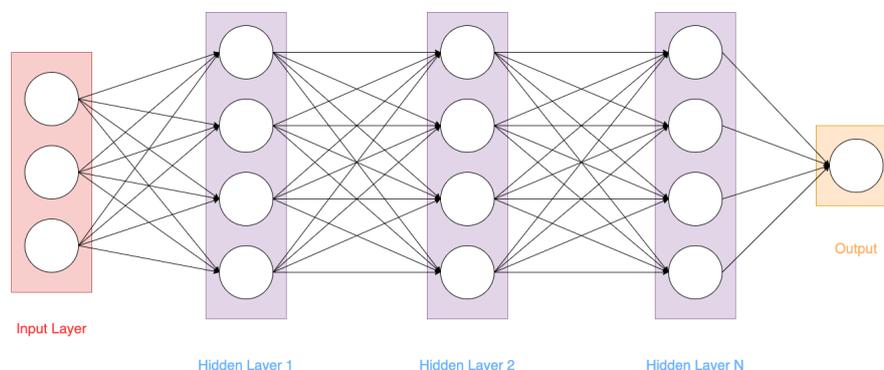
2 Kajian Pustaka

2.1 Studi Terkait

Pada [3] dilakukan penelitian berupa klasifikasi citra menggunakan CNN dengan menggunakan *dataset Caltech101*. Pada penelitian tersebut dilakukan praproses dan pengolahan pada data input yang berupa citra, praproses yang dilakukan yaitu *wrapping* dan *cropping*, selanjutnya citra juga diolah kembali dengan melakukan konversi menjadi grayscale dan merubah ukuran menjadi 140 x 140. Peneliti melakukan proses pelatihan dengan menggunakan tiga tahap, tahap pertama dilakukan proses *feedforward*, tahap kedua dilakukan proses *backpropagation* dan tahap ketiga melakukan perhitungan *gradient* dari jaringan konvolusi. Pengujian dilakukan dengan cara menggunakan bobot dan bias dari proses pelatihan dan data yang diuji adalah lima kategori unggas dan tiga kategori lainnya, hasil klasifikasi dari pengujian menunjukkan bahwa persentase keberhasilan pada kategori unggas sebesar 20% dan pada kategori lainnya menunjukkan persentase keberhasilan sebesar 50%, bila digabungkan klasifikasi citra menghasilkan persentase kebenaran 20% hingga 50%.

2.2 Dasar Teori

2.2.1 Convolution Neural Network (CNN) CNN merupakan versi peningkatan dari *Multi Layer Perceptron* (MLP), yang mana MLP memiliki satu dimensi pada setiap neuronnya sedangkan CNN melakukan peningkatan ukuran dimensi menjadi dua dimensi. CNN memiliki tiga layer utama, yaitu *Convolution Layer, Pooling Layer* dan *Fully Connected Layer*.



Gambar 1. Model MLP Sederhana

Pada jaringan MLP seperti gambar 1 memiliki beberapa *hidden layer* dengan *neurons* pada setiap layernya. Setiap hubungan antar *layer* memiliki parameter bobot yang dihasilkan. MLP menerima *input* data satu dimensi dengan *input* tersebut dilakukan penyebaran pada *hidden layer* hingga menghasilkan *output* [3].