BAB I
PENDAHULUAN

1. 1 Latar Belakang

Malware merupakan salah satu ancaman keamanan paling serius di internet saat ini, bahkan, masalah internet kebanyakan termasuk *junk mail* dan serangan *denial of service* memiliki *malware* sebagai penyebab utamanya[1], dalam keadaan ini sistem komputer yang terinfeksi oleh *malware* sering berlaku secara kolektif membentuk *botnet*, melakukan serangan jaringan dan banyak serangan lainnya. *Malware*, atau perangkat lunak berbahaya adalah istilah umum yang menggambarkan program atau *source code* yang berbahaya bagi sistem komputer[2] dimana program berbahaya ini membuatnya agresif dan sangat mengganggu, *malware* berusaha untuk menyerang, merusak, atau menonaktifkan sistem komputer, jaringan, tablet, dan perangkat *mobile*, bahkan mengambil kendali atas operasi perangkat.

Tak hanya berbahaya untuk sistem komputer saja *malware* bisa menjadi alat untuk merusak program aplikasi jika aplikasi tersebut tidak memiliki sistem keamanan yang tinggi, kerentanan dalam sistem aplikasi bisa menjadi celah *malware* untuk masuk kedalamnya dan merubah fungsi dari aplikasi itu sendiri. Saat ini, deteksi perangkat lunak berbahaya dilakukan terutama dengan metode berbasis heuristik dan *signature* yang berusaha untuk mengikuti evolusi *malware*. Metode berbasis *signature* telah banyak digunakan untuk perangkat lunak antivirus selama beberapa dekade[3]. Meskipun mengidentifikasi virus tertentu menguntungkan karena lebih cepat untuk mendeteksi keluarga virus melalui *signature* generik tetapi metode tersebut bisa dibilang tidak mampu untuk mengidentifikasi *malware* secara akurat dan keseluruhan jika *malware* yang menginfeksi adalah *malware* dengan evolusi terbaru yang belum pernah ada sebelumnya[4].

Ancaman *malware* telah menjadi semakin dinamis dan rumit, dan, sebagai akibatnya, strategi kecerdasan buatan harus lebih fokus untuk keamanan siber, karena mereka dianggap lebih ideal untuk mengatasi insiden *malware* modern[3].

Khususnya, jaringan saraf, dengan fungsi kinerja generalisasi yang kuat, mampu mengatasi sejumlah besar ancaman siber. Jaringan saraf, juga dikenal sebagai jaringan saraf tiruan (ANNs) atau jaringan saraf simulasi (SNN) adalah bagian dari *Machine Learning* dan merupakan jantung dari algoritma *Deep Learning*, nama dan struktur mereka terinspirasi oleh otak manusia, meniru cara neuron biologis saling memberi sinyal [5].

Convolutional Neural Networks (CNN) dianalogikan dengan ANN tradisional karena mereka terdiri dari neuron yang mengoptimalkan diri melalui pembelajaran. Setiap neuron masih akan menerima input dan melakukan operasi (seperti produk skalar diikuti oleh fungsi non-linier) -dasar dari ANN yang tak terhitung jumlahnya. Dari input vektor gambar mentah hingga output akhir skor kelas, seluruh jaringan masih akan mengekspresikan *single perceptive score function* (bobot). Lapisan terakhir akan berisi *loss funtion* yang terkait dengan kelas, dan semua tips dan trik reguler yang dikembangkan untuk ANN tradisional masih berlaku[6]. Dalam beberapa tahun terakhir metode pembelajaran mendalam seperti jaringan Saraf Konvolusional (CNN) dan Jaringan Saraf Berulang (RNN) digunakan di bidang keamanan informasi dan memberikan hasil yang lebih besar daripada metode lama[7]. Terutama penggunaan CNN telah terbukti alat yang lebih kuat. Jika kita dapat merubah berbagai jenis *malware* ke dalam gambar dan menjadikannya sebagai data input pada CNN untuk dikenali. Karena kemampuan mereka untuk fokus pada bagian lokal dari data input, jaringan saraf konvolusional sangat efektif dalam pengenalan dan klasifikasi gambar[8].

CNN juga telah terbukti efektif untuk jenis pemrosesan bahasa alami tertentu, yang memiliki implikasi untuk keamanan siber. Ini akan membantu administrator sistem untuk mengembangkan metode yang efisien untuk deteksi *malware*. Oleh karena itu penerapan sistem deteksi *malware* berbasis kecerdasan buatan menjadi pilihan yang tepat untuk menciptakan metode perlindungan dengan algoritma yang akan dapat mendeteksi dan menganalisis *malware* yang tidak diketahui dan dengan demikian tidak hanya meningkatkan keamanan tetapi juga menjadi metode cerdas untuk mendeteksi *malware* secara otomatis.

1. 2 Perumusan masalah

Berdasarkan latar belakang masalah yang di uraikan, dapat dirumuskan sebuah permasalahan yaitu bagaimana implementasi deteksi *malware packer* komputer menggunakan basis *Convolutional Neural Network* (CNN), alogaritma CNN akan mendeteksi file *packer* yang ada di komputer untuk mengenali apakah file tersebut aman atau sebuah *malware*.

1. 3 Tujuan Penelitian

Berdasarkan uraian diatas, maka peneliti memiliki tujuan yang dapat dicapai dari tugas akhir ini adalah sebagai berikut:

a. Implementasi *Convolutional Neural Network* untuk *mendeteksi malware packer*.
b. Memberikan hasil deteksi yang relevan dan analisa *Convolutional Neural Network* dengan akurasi tinggi.

1. 4 Batasan Masalah

Untuk mempersempit pembahasan pada skripsi ini, maka dibuat batasan-batasan sebagai berikut:

a. *Convolutional Neural Network* akan menjadi fokus utama.
b. Pembuatan *source code* menggunakan bahasa pemrograman Python dan dibantu oleh librari yang dibutuhkan.
c. Penulisan *source code* menggunakan Google Colab.
d. Dataset *malware* diperoleh dari Github.
e. Penelitian mengambil sampel *malware packer* berupa PE.

1. 5 Manfaat Penelitian

Penulisan laporan tugas akhir ini memiliki manfaat sebagai pengetahuan untuk menerapkan system kecerdasan buatan terhadap sebuah program untuk mendeteksi sebuah *malware* dan juga tugas akhir ini diharapkan bermanfaat untuk menambah pengetahuan mengenai system defensif sebuah program pada komputer untuk melindungi computer dari serangan *malware*.