EXPERT JURNAL MANAJEMEN SISTEM INFORMASI DAN TEKNOLOGI

Terakreditasi SINTA Peringkat 5
Surat Keputusan Menristek / Kepala BRIN No. 200/M/KPT/2020
Masa berlaku mulai Vol. 8 No. 1 tahun 2018 s.d. Vol. 12 No. 2 tahun 2022
Terbit online pada laman web jurnal http://jurnal.ubl.ac.id/index.php/expert

Peran *Big Data* dan *Deep Learning* untuk Menyelesaikan Permasalahan Secara Komprehensif

Novanto Yudistira
Fakultas Ilmu Komputer, Universitas Brawijaya
Jawa Timur, INDONESIA
yudistira@ub.ac.id

Abstrak – Peran sain data besar *(Big Data)* dan pembelajaran mesin dewasa ini tidak dapat terelakkan terutama untuk menganalisis data dan memberikan kecerdasan pada komputer agar bekerja secara otonom untuk menyelesaikan suatu pekerjaan tertentu. Perkembangan teknologi sensor dan internet membuat ketersediaan data tersebut melimpah yang selanjutnya dapat dilakukan analisis data dalam jumlah yang besar. Hal tersebut mempengaruhi bagaimana cara pandang komputasi dalam berbagai macam bidang baik ilmu alam maupun sosial. Data yang terkumpul dapat berupa beragam format dengan laju pertambahan yang cepat dan dinamis. Kita perlu algoritma atau model yang mumpuni untuk memahami dan menggali pengetahuan pada set data yang besar tersebut beserta rancangan modelnya yang secara otomatis mempunyai kemampuan memprediksi atau mendeteksi. *Deep Learning* dengan kapasitasnya yang besar serta hubungan korelasi antar neuron yang sangat banyak diharapkan mampu menjawab tantangan tersebut didukung oleh beberapa penelitian terkini pada penerapannnya di berbagai bidang keilmuan. Dalam *paper* ini akan dipaparkan contoh pemanfaatan *Deep Learning* pada *Big Data* yang telah kita lakukan pada pengenalan video aksi manusia pada Youtube, Segmentasi pada sel berskala besar, citra dada *x-ray* dan data *time-series* multi variabel hubungannya dengan pandemi COVID-19.

Kata Kunci: Big Data; Deep Learning; Permasalahan Komprehensif.

1. Pendahuluan

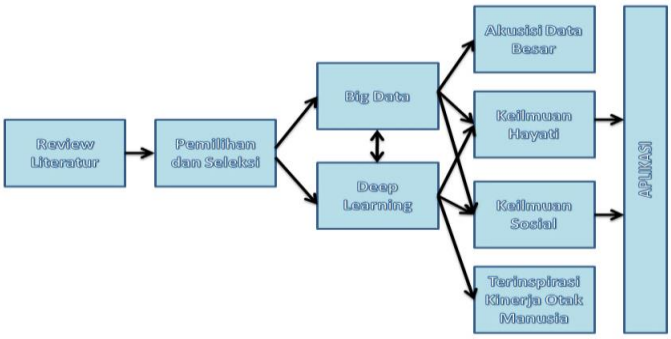
Kehadiran *Big Data* tidak terlepas dari kemajuan teknologi sensor, internet dan penyimpanan dimana data dapat diakuisisi secara otomatis tanpa mengenal ruang dan waktu. Melimpahnya data membutuhkan pemrosesan data yang tepat yang dapat digunakan untuk analisa data secara komprehensif maupun pelatihan pembelajaran mesin untuk keperluan kecerdasan buatan. Oleh karenanya, algoritma *Deep Learning (DL)* yang merupakan model algoritma terbaru untuk pembelajaran mesin mampu mengakomodir data yang melimpah oleh karena banyaknya parameter dan model pelatihan berbasiskan data yang mampu menangkap karakteristik data besar yang kaya. Beberapa aplikasi yang memanfaatkan model ini berkembang terutama pada bidang visi komputer, pencitraan medis, analisa teks, analisa data, prediksi dan lain sebagainya di berbagai dimensi keilmuan. Di Indonesia, seiring adaptasi infrastruktur jaringan internet yang semakin maju, meningkatnya ekonomi, dan berkembangnya perusahaan-perusahan teknologi informasi skala kecil, menengah, maupun besar tentunya kebutuhan akan analisa data yang masif menjadi prioritas. Analisa data besar dan pemanfaatannya dalam kecerdasan buatan tentunya akan berdampak pada bidang ilmu alam, sosial, medis dan lain sebagainya. Namun demikian, kita perlu memetakan hubungan *Big Data*, DL, dan pemanfaatannya di berbagai bidang mulai dari kehadiran teknologi sensor, karaketeristik DL dibanding pembelajaran mesin tradisional, sampai pada contoh-contoh aplikasi yang sudah dikembangkan dalam rentang 5 tahun belakang ini. Hal tersebut belum pernah dijelaskan secara runut berdasarkan literatur-literatur yang ada terutama dalam bahasa Indonesia. Berdasarkan hal tersebut maka penelitian ini mengumpulkan beberapa literatur yang dapat mengkonstruksi munculnya terminologi *Big Data* dewasa ini, menghubungkan antara *Big Data* dan DL, dan memberi gambaran aplikasi yang berkembang dan tren ke depannya secara komprehensif.

2. Metodologi

Dalam penelitian ini digunakan tahapan-tahapan mulai dari pemilihan referensi guna mendapatkan pemetaan mulai dari hubungan *Big Data* dan DL sampai pada aplikasi pada berbagai bidang secara komprehensif. Tahapan-tahapan tersebut digambarkan ini dilakukan dengan Gambar 1.

Sesuai yang ditunjukkan pada Gambar 1, tahap pertama dilakukan dengan melakukan studi literatur terhadap pustaka-pustaka yang berkaitan dengan *Big Data* dan DL mulai tahun 2009 sampai 2020. Pencarian artikel dimulai dari topik hulu dimana perangkat akuisisi data besar dan permulaan dan dinamika kemelimpahan dibahas. Selain itu, artikel-artikel tentang pemrosesan data besar pada berbagai bidang juga dicari dan dipilih untuk kita kategorikan berdasarkan bidang yang diselesaikannya. Pada pencarian artikel tentang DL, kita juga membahas tentang bagaimana DL ini terinspirasi dari fenomena biologis bada otak manusia hingga dapat ditiru secara matematis. Kemudian, kita cari ketersambungan antara DL dan *Big Data* sehingga dapat ditangkap dan diceritakan keterkaitannya serta potensi ke depannya

dalam secara aplikatif. Aplikasi DL dan *Big Data* tentunya sedang berkembang dan kita cari artikel-artikel terkini yang sedang meneliti tentang aplikasi yang memanfaatkan data yang banyak dan DL pada berbagai topik sebagai contoh tren aplikasinya. Kita memilih pada studi kasus yang umum yaitu klasifikasi citra umum, citra medis, hingga prediksi dengan data berskala besar.



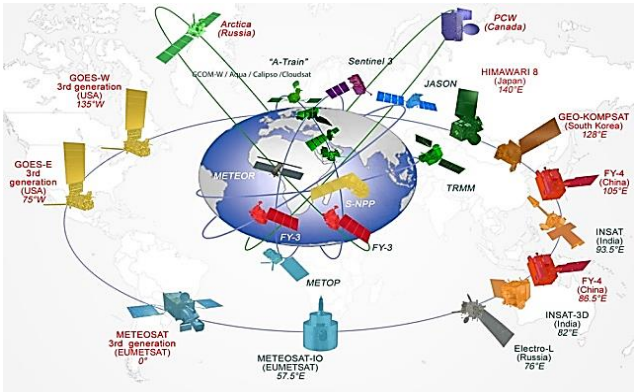
Gambar 1. Diagarm Alur Penelitian

3. Hasil dan Pembahasan

A. Peran Teknologi Sensor dan Teknologi Internet dalam Akuisisi Data Besar

Beberapa tahun terakhir kita menjadi saksi berkembangnya miliaran sensor, perangkat bergerak, kamera, dan lain sebagainya yang terhubung ke dunia maya baik itu jaringan internet maupun basis data. Setiap hari sensor-sensor pintar ini secara otonom menangkap dan kemudian merekam data secara besar dengan laju yang cepat sampai seukuran Zetta Bytes. Seiring infrastruktur pengumpulan data yang semakin canggih, kita juga menghadapi peluang dan tantangan baru [1]. Dengan penginderaan dan pemrosesan data multimedia, bersama dengan alokasi sumber daya, pengoptimalan kualitas layanan (QoS), keamanan dan privasi, platform, alat, dll., analisa *Big Data* telah menjadi instrumen utama dan penting utamanya untuk IoT yang mampu mengumpulkan dan mengolah data besar didukung infrastruktur yang mumpuni [2]. Teknologi dan paradigma baru yang muncul termasuk komputasi awan, pembelajaran mendalam (DL), virtualisasi fungsi jaringan, penginderaan oleh banyak perangkat seluler, dan jaringan 4G maupun yang baru 5G sangat dibutuhkan untuk memainkan peran dalam era baru sekarang ini.

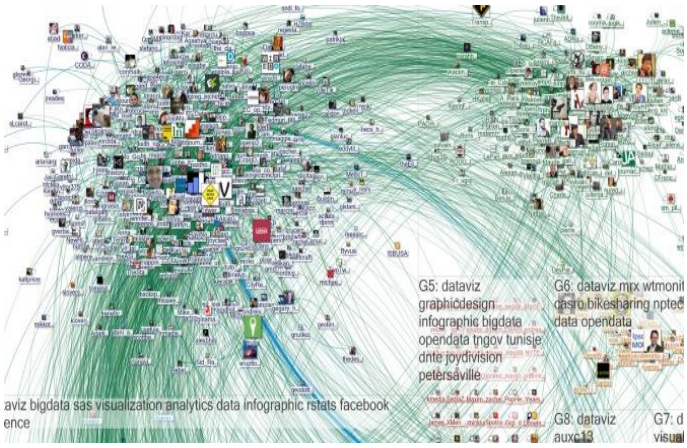
Big Data adalah istilah ekspansif untuk cakupan informasi yang begitu luas atau kompleks sehingga aplikasi penanganan informasi biasa tidak memadai [3]. Data besar dihadirkan ke dunia dengan pendekatan pemrosesan digital dimana saja tanpa mengenal batasan ruang dan waktu. Data-data itu dapat terkumpul secara crowdsourcing melalui media teknologi internet dengan disadari ataupun tidak seperti penggunaan perangkat bergerak hingga satelit (Gambar 2). Crowdsourcing berasal dari berbagai sumber yang beragam dan terkumpul hingga menjadi informasi yang sangat besar. Kesulitan dan tantangan yang akan dihadapi dalam pemrosesan informasi yang besar adalah investigasi data, proses akuisisi data, teknik kurasi informasi, teknik visualisasi, berbagi data, penyimpanan data, pertukaran data, persepsi informasi, hingga keamanan data itu sendiri. Informasi yang sangat besar sangat penting karena memungkinkan kita untuk mengumpulkan, menyimpan, mengawasi, dan mengontrol informasi yang sangat banyak dengan kecepatan yang benar pada waktu yang tepat untuk mengambil bit pengetahuan yang sesuai. Dengan proses akuisisi *Big Data* yang benar memungkinkan kita untuk memvirtualisasikan dan menyimpan informasi secara efektif melalui repositori berbasis awan dengan biaya paling memadai.



Gambar 2. Akuisisi Data dari Beberapa Satelit

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 79

Beberapa hal yang membuat dinamisnya *Big Data* [4] adalah dari sisi Volume (Skala): Berapa banyak peningkatan informasi yang ada semisal meningkatnya volume data 44 kali dalam rentang 2009 sampai 2020 mulai dari ukuran 0.8 zetta byte hingga 35 zetta byte [5]. Hal ini berarti volume data meningkat secara eksponensial dari tahun ke tahun. Hal kedua yang penting untuk mesifati *Big Data* ini adalah *velocity* (kecepatan) yaitu seberapa cepat informasi tersebut berubah atau diakuisisi. Data yang diakuisisi dengan cepat maka dinamikanya semakin tinggi sehingga perlu diproses dengan cepat pula dan adaptif. Dari hal tersebut timbullah teknik analisis data online sebagai contohnya, e-promosi dimana berdasarkan lokasi kita saat ini, riwayat pembelian kita dan apa yang kita suka dapat dianalisis sehingga mampu mengirimkan promosi yang tepat sasaran dan tepat guna pada saat itu juga. Contoh lainnya adalah dalam pemantauan perawatan kesehatan semisal sensor yang memantau aktivitas dan tubuh anda sehinggga hasil prediksi kondisi abnormal dapat ditangani dengan segera. Sifat yang ketiga adalah varietas (kompleksitas) yaitu ragam tipe informasinya. Beragam format baik itu tipe atau strukturnya dapat menjadi satu tantangan dalam pemrosesan data. Lebih spesifik, ragam dapat berupa teks, numerik, gambar, audio, video, urutan, deret waktu, data media sosial, dll. Dilihat dari model akuisisi datanya, apakah data itu statis atau streaming [6] dalam suatu aplikasi dapat mempengaruhi ragam data yang terkumpulkan.



Gambar 3. Visualisasi hubungan antar entitas dalam sosial media (Sumber gambar: NodeXL Twitter Search oleh Marc Smith [7])

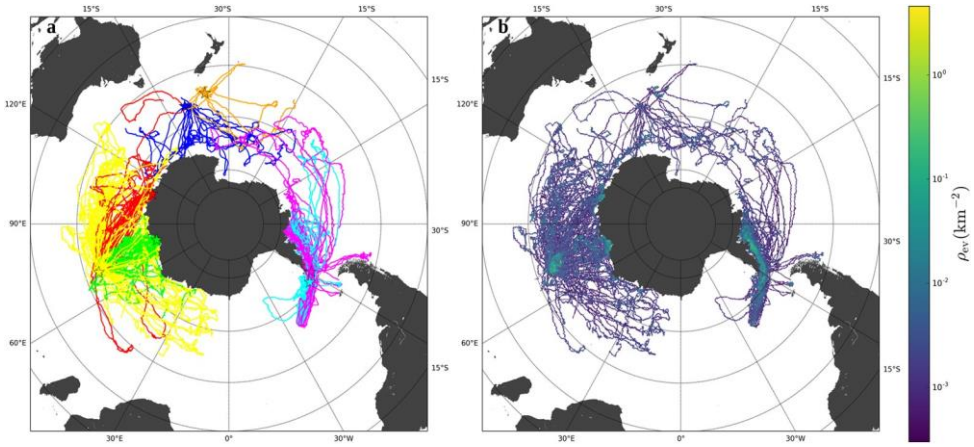
B. Analisis Data Media Sosial Berskala Besar

Analisis *Big Data* baru-baru ini muncul sebagai area penelitian yang penting oleh karena popularitas Internet dan munculnya teknologi web 2.0. Selain itu, perkembangan dan adopsi aplikasi media sosial telah memberikan peluang dan tantangan yang luas bagi para peneliti dan praktisi. Sejumlah besar data yang dihasilkan oleh pengguna yang menggunakan platform media sosial adalah hasil integrasi informasi yang detail tentang latar belakang dan aktivitas sehari-hari mereka (Gambar 3). Data bervolume besar yang dihasilkan yang dikenal sebagai *Big Data* telah diteliti secara intensif baru-baru ini. Berdasarkan tinjauan karya-karya terbaru, mereka berusaha menyajikan perspektif yang luas tentang topik penelitian analitik *Big Data* media sosial untuk berbagai bidang. Beberapa literatur dapat dikelompokkan dalam bidang ilmu alam, sosial, ekonomi, politik, dan lain sebagainya. Tidak hanya itu, studi-studi yang ada juga membandingkan kemungkinan teknik analisis data besar beserta atribut dan kualitasnya. Permasalahan dalam penelitian menggunakan analisis data besar di berbagai bidang ini sangat menantang untuk diteliti secara komprehensif dan multi-dimensi dengan melibatkan banyak pihak dan pakar.

C. Meningkatkan Penelitian dan Penemuan dalam Ilmu Hayati

Dengan data besar dapat membantu mempercepat pertumbuhan dan penemuan di berbagai industri, organisasi dan keilmuan hayati. Langkah-langkah besar telah dilakukan dengan mendorong pengumpulan secara otomatis melalui sensor kecerdasan untuk mengakuisisi data pada berbagai repositori. Berbagai macam platform analitik data besar telah tersedia, baik secara komersial maupun *open source*. Namun, tantangan utama bagi kenayakan organisasi adalah kurangnya keahlian yang mendalam untuk mendukung inisiasi analisis data besar. Bagaimana kita dapat memperoleh dan menyiapkan data secara efisien? Bagaimana kita dapat mengetahui metode pengolahan secara praktis dan cerdas dari sejumlah besar data yang telah kita kumpulkan? Tentunya merupakan tantangan tersendiri.

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 80



Gambar 4. *Big Data* dengan menelurusri pola pergerakan gajah laut pada rentang waktu tertentu menggunakan teknologi sensor [8]

Sejumlah besar data sedang dihasilkan baik dalam keilmuan biologi, ilmu alam, industri dan institusi perawatan kesehatan, yang menjanjikan untuk memajukan pemahaman kita tentang berbagai sistem dan penyakit biologis, pengembangan biokatalis dan obat-obatan baru, jamu tradisional serta pemberian perawatan pasien dan pengurangan yang lebih efiseien dalam segi biaya, dll. Contohnya adalah analisis pola pergerakan pada skala populasi dan bahkan spesies gajah laut memanfaatkan *Big Data* (Gambar 4). Meningkatnya jumlah data secara besar dengan sensor untuk pelacakan hewan memberikan peluang untuk untuk menganalisis data pelacakan dari 272 gajah laut (*Mirounga leonina*) di Samudera Selatan.

D. Meningkatkan Penelitian dan Penemuan dalam Ilmu Sosial

Apa arti *Big Data* bagi ilmu sosial kontemporer? Bagaimana kecepatan, variasi, dan volume aliran *Big Data* dapat digunakan untuk mendapatkan pemahaman yang lebih baik tentang fakta sosio-ekonomi yang kompleks? Apakah *Big Data* merupakan alat yang layak untuk mengatasi masalah sosial? Ketika data menjadi semakin berharga, siapa yang akan memiliki dan mengontrol akses ke data tersebut? Tentunya beberapa tantangan permasalahan yang perlu dijawab di masa depan.

Seiring pesatnya peningkatan jumlah data sosial yang dihasilkan dan yang tersedia, tren baru-baru ini bagi para peneliti dari ilmu sosial adalah memahami potensi *Big Data* dalam melengkapi metode penelitian tradisional dan nilainya dalam membuat keputusan. Memang, *Big Data* membutuhkan peninjauan kembali terkait teknik analisis data dengan cara mendasar di semua tahap mulai dari akuisisi dan penyimpanan data hingga transformasi dan interpretasi data. Secara khusus, tugas mengumpulkan dan menganalisis data yang merupakan inti dari alur analisa *Big Data* mempunyai tantangan yang mendesak dalam domain ilmu sosial. Jenis data yang tersedia terbagi dalam berbagai kategori: data sosial (misal Twitter feed, Facebook like), data tentang mobilitas dan lokasi geospasial (mis., data sensor yang dikumpulkan melalui ponsel atau citra satelit), data yang dikumpulkan dari berbagai sumber administrasi pemerintah dan multi bahasa. Selain itu, data sering kali terpecah-pecah di banyak sumber dan sering kali memerlukan terjemahan dari satu bahasa (atau format tertentu) ke bahasa lain dan, dalam beberapa kasus ekstrem, diperlukan komunikasi antar disiplin ilmu yang berbeda.

Beberapa masalah utama harus diselidiki dengan cermat seputar *Big Data* dalam ilmu sosial. Pertama, data yang hilang menjadi perhatian utama para peneliti ilmu sosial, terutama bagi mereka yang bertujuan untuk mempelajari efektivitas pendekatan berbasis data dalam proses pengambilan keputusan [9]. Kedua, data sosial yang dihasilkan dari interaksi dengan manusia seringkali tidak dapat diandalkan atau tidak valid. Oleh karena itu, proses pengumpulan data harus menggabungkan mekanisme untuk menemukan potensi ketidakakuratan dan mengukur sejauh mana ketidakakuratan tersebut tercermin dalam hasil tugas analisis data. Terakhir, kecepatan data sosial yang dihasilkan dari interaksi manusia melalui peningkatan sejumlah platform dan banyaknya perangkat yang saling berinteraksi menimbulkan beberapa tantangan berupa respon secara real-time dan efektif.

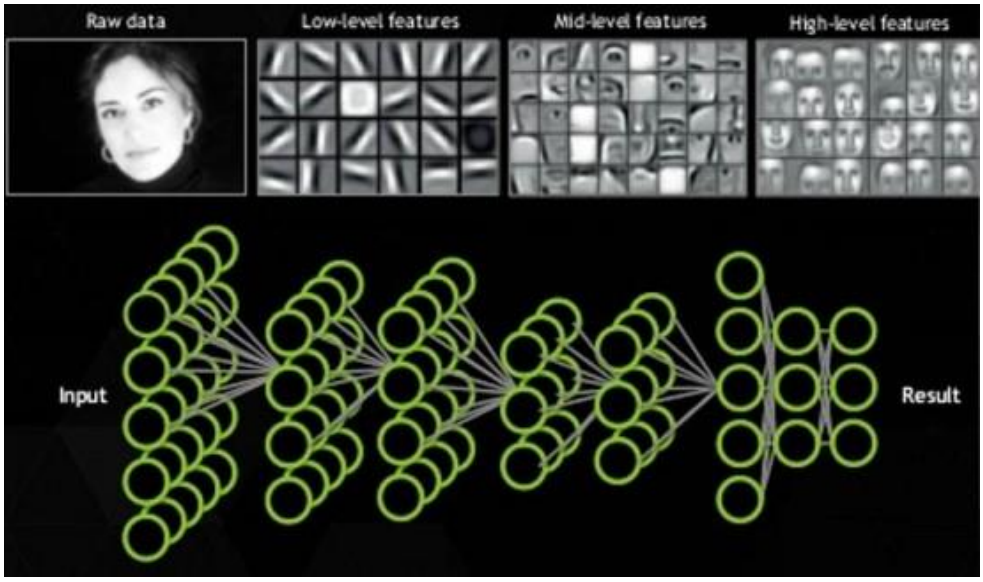
E. Prediksi *Big Data* Menggunakan *Deep Learning* yang Meniru Cara Kerja Otak Manusia

Pembelajaran mendalam (DL) memungkinkan model kecerdasan buatan (AI) untuk belajar dengan cara yang sangat mirip dengan cara manusia belajar, yaitu melalui pengalaman dan persepsi secara berkelanjutan [10]. Selayaknya kita mengajari bayi cara mengenali anjing dan kucing dengan menunjukkan banyak gambar dan hewan peliharaan nyata kepada mereka, kita dapat mengajari model AI berbasis pembelajaran mendalam cara mengenali gambar dan pola dengan memberikan banyak data kepada model.

Pembelajaran mendalam (DL) adalah metode kecerdasan buatan (AI) yang secara matematis meniru cara kerja otak untuk menangkap pola penting dari data yang besar. Dengan meniru operasi otak, sebuah model matematika pembelajaran mendalam yang ditulis dalam program komputer dapat menyaingi, atau bahkan mengungguli manusia dalam sejumlah fungsi, seperti

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 81

pengenalan gambar [11], kontrol motorik [12], dan pengenalan suara [13]. Selain itu, jaringan dalam dapat mengembangkan representasi yang lebih cocok dengan rekaman di neokorteks manusia atau primata non-manusia daripada model yang ada dalam ilmu saraf [14]. Hal ini menunjukkan bahwa pembelajaran mendalam (DL) menangkap sistem yang penting tentang cara kerja otak kita sendiri. Contoh arsitktur DL ini adalah Convolutional Neural Network (CNN).



Gambar 5. Model *Deep Learning* dengan jaringan syaraf tiruan dalam (Sumber gambar: NVIDIA)

Kunci dari pembelajaran "mendalam" (sebagai lawan dari pembelajaran "dangkal") adalah penggunaan jaringan *multilayer* yang memiliki "lapisan tersembunyi" antara masukan sensorik dan keluaran berupa respon (jaringan dangkal tidak memiliki lapisan tersembunyi). Namun, pembelajaran dengan banyak lapisan memerlukan algoritma matematis yang efisien untuk dapat melakukan inferensi. Untuk melatih jaringan neuron secara efektif, setiap neuron harus menerima bobot atas kontribusinya terhadap prediksi selama pelatihan sistem berlangsung. Dalam jaringan yang dangkal dengan jumlah neuron yang sedikit tentu mudah, karena setiap neuron mendorong perilaku atau hanya terletak satu koneksi sinaptik. Namun, dalam jaringan dalam dengan lapisan tersembunyi, pemberian bobot pada setiap neuron bergantung pada arsitektur jaringan.

Sebelum model algoritma kecerdasan DL mendapat perhatian, terdapat banyak prediksi data yang menggunakan ciri buatan untuk dimasukkan ke dalam algoritma pembelajaran mesin seperti *Support Vector Machine, Naive Bayes, Random Forest*, dll. DL semakin populer sejak kompetisi *Imagenet* yang diadakan oleh Li Fei Fei. Yaitu kompetisi prediksi data citra skala besar sekitar 14 juta data dan 1000 kategori [15]. Perubahan popularitas mengarah pada perubahan cara penanganan dan pembuatan ciri yang berguna bagi model untuk memprediksi data. DL yang diusulkan oleh Hinton [16] mencapai akurasi terbaik, dan juga dilihat lebih dalam pada setiap lapisan konvolusi hasil pelatihan (Gambar 5). Gambar tersebut menunjukkan jaringan menangkap pola lokal ke global dan lebih kontekstual seiring arsitektur lebih kedalam. Arsitektur DL yang sering digunakan untuk prediksi data citra adalah CNN (*Convolutional Neural Network*) seperti pada Gambar 5 yaitu serangkaian konvolusi melalui operasi dot product mulai dari input matriks yg besar dikonvolusi oleh filter matriks yg lebih kecil sehingga menghasilkan *feature map* dan setiap lapisannya dimasukkan dalam fungsi aktivasi *RectifiedLinear Unit* (ReLU) yg berfungsi menonjolkan atau mengeksitasi fitur-fitur yg menonjol dari hasil yg dikonvolusi oleh filter-filter tersebut untuk menghasilkan *activation map*. Sedangkan bagian *pooling* digunakan untuk mereduksi ukuran *activation map* tersebut. Akhirnya, pemrosesan dilanjutkan oleh lapisan penklasifikasi *(fc layers)* untuk dilakukan pendeteksian. Proses ini mirip dengan proses di otak dimana menemukan detail-detail kecil seperti *blob*, garis atau tepi berlanjut sampai bentuk yg lebih besar, abstrak dan global misal wajah.

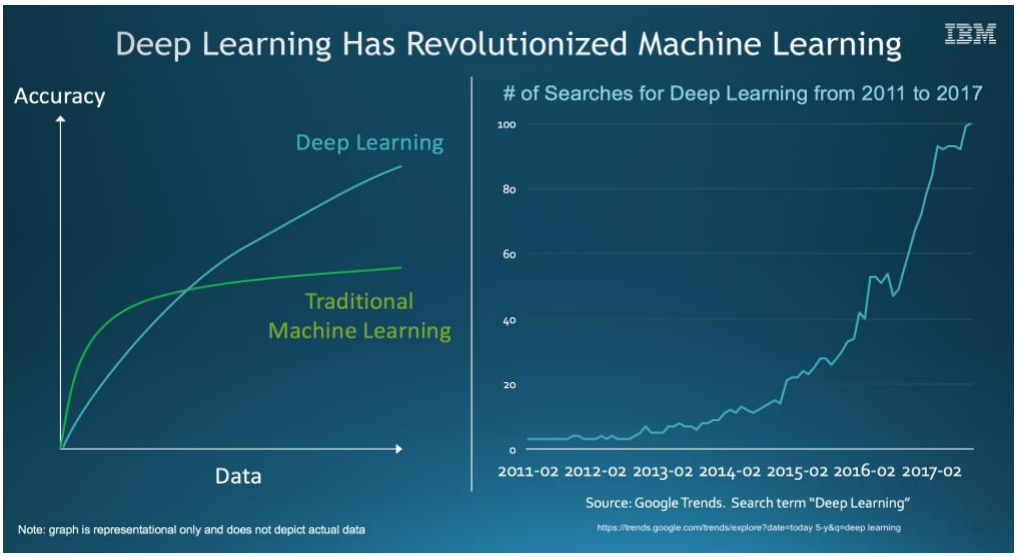
Karena banyaknya parameter, DL terkadang disebut *over-parameterized*, menyebabkan masalah lain untuk diselesaikan seperti *overfitting* apabila jumlah data latih tidak mencukupi. Cara mengatasinya biasanya dengan pemangkasan parameter menggunakan beberapa metode optimasi untuk menghasilkan arsitektur DL yang efisien. Aplikasi ini diterapkan pada berbagai topik, seperti visi komputer, pemrosesan bahasa alami, pengenalan suara, dll. Bahkan perkembangan terbaru telah mencapai sintesis gambar atau video membuat DL secara meyakinkan berkontribusi pada pengembangan pengenalan pola pada data besar ke tingkat berikutnya.

F. Mengapa *Big Data* Cocok dengan *Deep Learning* (Data-Driven)

Pembelajaran mendalam adalah metode pembelajaran mesin baru berdasarkan jaringan saraf yang belajar dan menjadi lebih akurat dengan memberi lebih

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 82

banyak data kepada model. Prinsip pembelajaran mendalam yang diterima secara luas ditampilkan di sisi kiri diagram pada Gambar 6 tersebut: model AI berbasis pembelajaran dalam memiliki akurasi yang jauh lebih tinggi daripada metode pembelajaran komputer tradisional, tetapi membutuhkan lebih banyak data untuk dilatih guna mencapai keakuratan tersebut. Di sebelah kanan, merupakan hasil penelusuran *Google Trends* tentang "*Deep Learning*" untuk menunjukkan bagaimana orang mencari lebih banyak informasi tentang pembelajaran mendalam (DL) selama beberapa tahun terakhir (Gambar 6).



Gambar 6. Perbedaan antar machine learning tradisional dan *Deep Learning*, dimana semakin banyak data, *Deep Learning* semakin unggul [17]

Pembelajaran mendalam memiliki banyak kegunaan di setiap industri, mulai dari analisis ritel, video *drone* hingga pencitraan medis untuk membantu dokter melakukan diagnosis. Bisnis dewasa ini dapat menggunakan jenis metode pembelajaran mesin canggih ini untuk mengekstrak pengetahuan dari data yang telah dikumpulkan dalam ruang penyimpanan yang besar dalam beberapa tahun terakhir. Sebagai contoh semisal bank yang saat ini sebagian besar menggunakan sistem berbasis aturan untuk deteksi penipuan, di mana aturan tersebut dapat menentukan serangkaian kondisi yang akan memicu peringatan tentang adanya penipuan.

Sebaliknya, mereka dapat menggunakan data penggunaan kartu kredit beberapa tahun terakhir untuk melatih model pembelajaran mendalam dan mempelajari lebih banyak data yang kita berikan. Faktanya, bahkan setelah kita menerapkan model AI ke dalam produksi, model ini dapat terus belajar dari jutaan transaksi kartu kredit yang diproses oleh bank setiap hari. Keuntungan dari pendekatan ini adalah bahwa model AI secara otomatis mempelajari situasi kecurangan baru berdasarkan pengalaman yang terjadi, daripada mengandalkan seorang ilmuwan data yang menulis aturan baru untuk setiap jenis situasi yang berbeda.

Dua hal utama yang membuat DNN menjadi pilihan utama, yaitu daya komputasi dan *Big Data*. Untuk menggunakan metode AI ini, sebuah bisnis perlu memproses data dalam jumlah besar, dan pendekatan tersebut memerlukan infrastruktur teknologi informasi (TI) yang sesuai dengan tugasnya. Daya komputasi tinggi seperti *Graphical Processing Unit* (GPU) dapat mengatasi keterbatasan daya komputasi dibandingkan apabila hanya menggunakan *Central Processing Unit* (CPU). DNN modern memiliki berjuta parameter untuk dioptimalkan, yang menuntut memori komputasi tinggi untuk seluruh *epoch* atau iterasi. Selain itu, DNN juga harus disuplai dengan *Big Data* secara berulang. Bukti kebutuhan *Big Data* ditunjukkan dengan penggunaan repositori data citra, ImageNet, yang memuat 14 juta citra dan 1000 kelas untuk melakukan klasifikasi citra. Untuk segmentasi objek, set data MS-COCO menjadi tolok ukur yang populer. Ada juga kumpulan data dalam klasifikasi video aksi serta meringkas teks atau dokumen. *Dataset* sentiment40 digunakan sebagai pra-pelatihan untuk tugas analisis sentimen. Ulasan WordNet dan Yelp digunakan dalam tugas Pemrosesan Bahasa Alami (NLP). *Dataset* dengan berjuta lagu merupakan pilihan yang tepat untuk digunakan sebagai bahan pelatihan pengenalan lagu atau set data Urban Sound yang digunakan untuk mempelajari berbagai sinyal yang dihasilkan oleh suara oleh berbagai sumber. Saat ini ada *dataset Google object detection* yang berisi 15000 video dan 4 juta gambar merupakan salah satu yang populer dalam tugas deteksi objek.

G. Contoh Data Besar Berupa Det Data Citra dan Video Berskala Besar

Hampir setiap hari ditemukan terobosan dalam pembelajaran mesin dalam tidak hanya terbatas pada teks atu dokumen tetapi juga citra gambar. Contoh aplikasi dari pengenalan gambar adalah *Google Image*, pengenalan wajah, pengenalan manusia dalam CCTV, dll. Untuk menyempurnakan algoritme DL dalam mengenali dan memprediksi pola dalam data, Kita perlu memberi mereka sejumlah besar data citra yang sudah diberi tag

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 83

untuk diuji dan dipelajari oleh algoritme. Terdapat beberapa set data besar citra seperti Open Images, YouTube8-M, dan Imagenet yang menyediakan jutaan citra beranotasi yang berguna bagi peneliti untuk melatih model algoritma mereka. ImageNet adalah kumpulan data gambar atau citra yang diatur menurut hierarki WordNet. Dalam set data WordNet, beberapa kata atau frasa kata disebut "set sinonim" atau "synset" pada setiap kategorinya. Ada lebih dari 100.000 synset di WordNet, kebanyakan dari mereka adalah kata benda (80.000+). Di ImageNet, yang merupakan set data citra besar masing-masing kategori berisi sekitar 1000 gambar (Gambar 7). Gambar dari setiap konsep dikontrol kualitasnya dan diberi label oleh manusia. Secara bertahap Imagenet akan menyediakan sekitar puluhan juta gambar yang diurutkan dengan rapi mengikuti konsep dalam hierarki WordNet.

Gambar 7. Set data citra skala besar yang digunakan untuk pelatihan *Deep Learning* [15]

Proyek ImageNet terinspirasi oleh permasalahan yang berkembang di bidang penelitian gambar dan visi yaitu kebutuhan akan lebih banyak data. Sejak dimulainya era digital dan adanya transaksi data skala web, para peneliti di bidang ini telah bekerja keras untuk merancang algoritme yang semakin canggih untuk mengindeks, mengambil, mengatur, dan membuat anotasi data multimedia. Tetapi penelitian yang baik membutuhkan sumber daya yang baik puka. Untuk mengatasi masalah ini, set data citra dalam skala besar (koleksi pribadi gambar digital, atau video kita, atau database mesin telusur web komersial), akan sangat membantu para peneliti dan praktisi di bidang kecerdasan buatan (AI) pada visi komputer. Hal tersebut memotivasi kalangan peneliti dan akademisi AI visi komputer untuk membuat ImageNet.

Gambar 8. Set data video berskala besar dari Youtube [18]

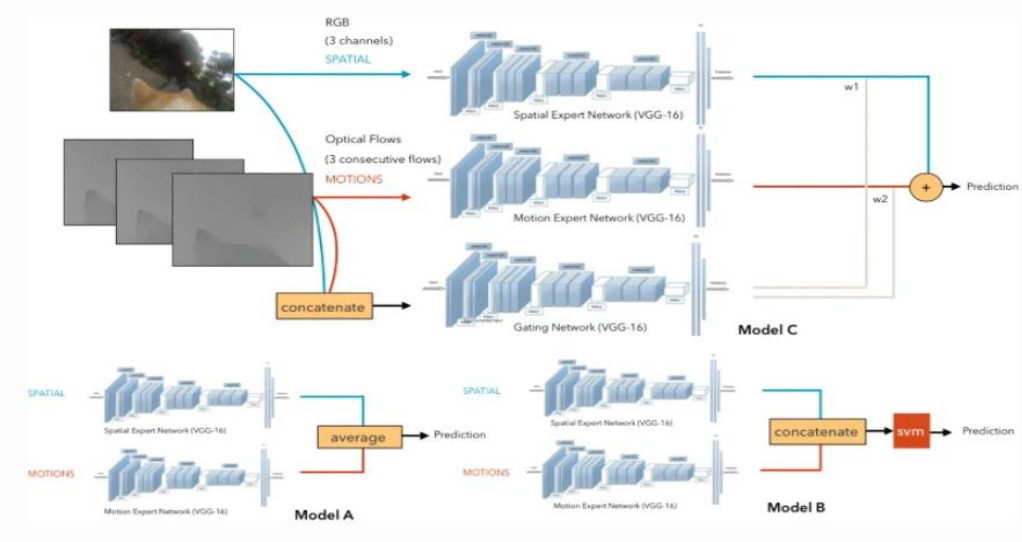
Gambar dan video telah ada di mana-mana di internet (Gambar 8), yang mendorong pengembangan algoritma yang dapat menganalisis konten semantiknya untuk berbagai aplikasi pengenalan termasuk penelusuran

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 84

dan peringkasan dengan beberapa aplikasi yang terkenal seperti Youtube, Vimeo, ataupun video yang diambil dari kamera ataupun GoPro. Baru-baru ini, algoritma *Convolutional Neural Networks* (CNN) [19] yang merupakan salah satu arsitektur DL yang telah terbukti sebagai model algoritma yang efektif untuk memahami konten gambar, memberikan hasil yang canggih pengenalan gambar yang canggih, melakukan segmentasi citra, serta deteksi objek. Faktor pendukung utama di balik kesuksesan ini adalah set data berlabel yang besar (*Big Data*) dan algoritma CNN yang mempunyai jaringan dengan jutaan parameter dan mendukung proses pembelajaran pada komputer. Dalam kondisi ini, CNN telah terbukti kuat dalam belajar dan dapat ditafsirkan melalui fitur gambar pada setiap lapisannya (Gambar 5). Didorong oleh hasil positif dalam domain citra ini, kinerja CNN di klasifikasi video skala besar diinvestigasi di mana jaringan memiliki akses tidak hanya ke citra tunggal berupa gambar statis tetapi juga evolusi temporal yang kompleks.

H. Studi Kasus: Kecerdasan Buatan Mendeteksi Jenis Aksi dari Video

Manusia dengan mudah mengenali dan mengidentifikasi tindakan aksi manusia dalam video, akan tetapi pada komputer otomatisasi merupakan hal yang menantang oleh karena variasi dan kompleksitasnya tentunya membutuhkan data yang besar untuk pelatihan modelnya. Aplikasi yang terkenal untuk dapat mendeteksi dan menklasifikasikan video adalah pada algoritma Youtube. Pengenalan aksi manusia dalam video sangat menarik untuk aplikasi seperti pengawasan otomatis, pemantauan perilaku lansia, interaksi manusia-komputer, pengambilan video berbasis konten, dan ringkasan video. Dalam memantau aktivitas kehidupan sehari-hari lansia, misalnya, pengenalan tindakan atomik seperti "berjalan", "membungkuk", dan "jatuh" dengan sendirinya sangat penting untuk analisis aktivitas.



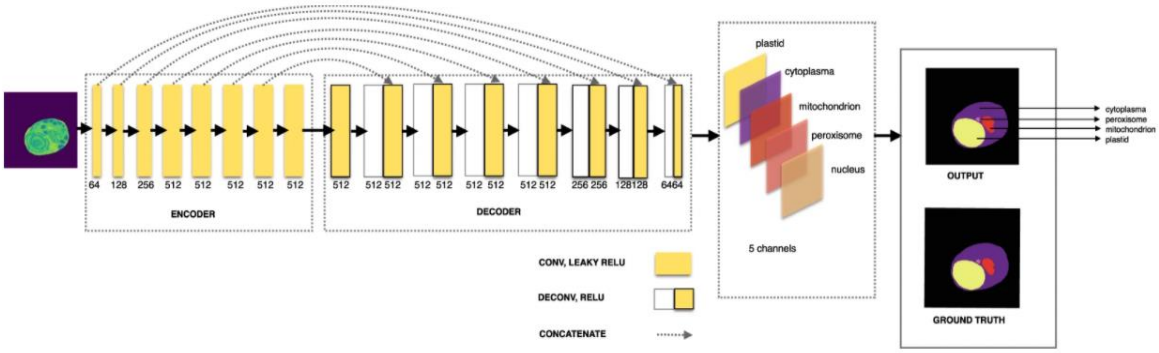
Gambar 9. Model *Deep Learning* dalam mendeteksi aksi manusia dalam video [20]

Kita mengajukan sebuah model DL yang mempu manangkap pola pada data yang besar benama *Gating CNN* yang merupakan model DL yang memanfaatkan satu jaringan CNN yang berlaku sebagai *associative cortex* yang mempu mengatur bobot jaringan CNN lainnya secara otomatis [20]. Perbandingan dengan metode *state-of-the-art* menggunakan set data UCF-101 (Gambar 8) berisi sekitar 13000 video dengan transfer bobot dari set data ImageNet, menunjukkan bahwa *Gating CNN* memberikan akurasi tertinggi dibanding eksperimen dengan model algoritma lainnya.

I. Studi Kasus: Kecerdasan Buatan Untuk Melakukan Segmentasi Sel ke dalam Beberapa Jenis Organel

Interpretasi sel hidup yang ditangkap dalam waktu lama menggunakan mikroskop elektron pemindaian canggih (SEM) berguna untuk mengidentifikasi perilaku sel sekuensial dan perubahan lokal berdasarkan lingkungannya (Golding, C., 2016) [21]. Namun, dalam beberapa tahapan organel dalam sel alga merah uniseluler primitif Cyanidioschyzon merolae (C. merolae) selama pembelahan sel akan muncul dalam berbagai posisi dan waktu. Kemajuan teknologi mikroskop dan pencitraan dengan resolusi spasial dan temporal sangat membantu untuk menentukan lokasi pola sel untuk menilai informasi sel biologisnya. Dengan demikian evaluasi objektif akan posisi kontur sel yang akurat akan bermanfaat untuk memantau siklus sel secara sekuensial selama mitosis berlangsung. Selain itu, beberapa metode otomatis telah dikembangkan untuk menganalisis peristiwa mitosis dan morfologinya dalam berbagai studi patologi.

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 85



Gambar 10. Model *Deep Learning* dalam melakukan segmentasi sel (Sumber: [22])

Kami mengajukan suatu metode DL menggunakan fungsi *loss* yang baru sehingga dapat mengatasi dan mempelajari data yang tidak seimbang dari sejumlah sekitar 3000-an data [22]. Data yang tidak seimbang tersebut terjadi karena ada beberapa organel yang frekuensi kemunculannya sangat jarang dan kecil sehingga secara kuantitas terlampau minim. Berdasarkan hasil eksperimen dan validasi, model yang kami ajukan menunjukkan kehandalan dan efisiensi pendekatan terutama dalam membedakan pola seluler yang tidak jelas dan frekuensi kemunculannya rendah yang pada akhirnya membantu ahli biologi dalam membuat keputusan yang handal selama proses *mitosis* untuk memahami fungsi dan perilaku sel sekuensial.

J. Studi Kasus: Pandemi COVID-19

Penyakit COVID-19 (Coronavirus Disease 2019) pertama kali diidentifikasi di provinsi Hubei Cina melalui adanya laporan jenis Pneumonia yang tidak diketahui penyebabnya. Semenjak 31 Desember 2019, COVID-19, yang mana virus tersebut bernama asli SARS-CoV-2, telah menyebar cepat hingga menjadi sebuah pandemi baru [23][24] (WHO,2020). Jenis virus baru ini mulai menyebar dari Wuhan ke sebagian besar Cina dalam rentang 30 hari [25] meskipun penyebarannya itu tidak sesignifikan dibanding yang terjadi di dalam provinsi Hubei itu sendiri. Sedangkan di Amerika Serikat, tujuh kasus pertama dilaporkan terjadi pada tanggal 20 Januari 2020 hingga mencapai lebih dari 300.000 kasus pada tanggal 5 April 2020 [26]. Virus ini tidak hanya dapat menular antar manusia akan tetapi juga melalui udara [27]. Termasuk virus jenis ini adalah Virus SARS-CoV dan MERS-CoV yang dapat menyebabkan sindrom pernafasan akut parah hingga dapat menyebabkan kematian pada manusia.

Pemanfaatan kecerdasan buatan akan membantu mengurangi dampak dari kurangnya alat tes RT-PCR sehingga meminimalisir biaya dan waktu tunggu pengujian. Sedangkan citra radiologi telah banyak digunakan dalam pencitraan medis sehingga dapat bermanfaat pula dalam mendeteksi COVID-19. Para peneliti menyatakan bahwa penggabungan fitur citra klinis dengan hasil laboratorium dapat membantu dalam deteksi dini COVID-19 [28]. Citra radiologi yang diperoleh pasien dengan kasus COVID-19 terdapat informasi yang berguna untuk diagnosa.

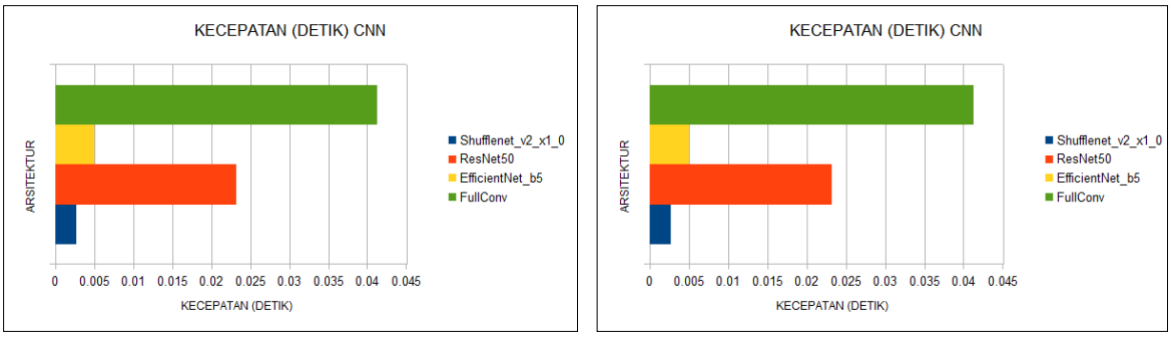
Pada penelitian ini, dataset citra sinar-X diperoleh dari sumber berbeda yang merupakan hasil dari diagnosis COVID-19 yaitu dari Cohen JP [29] dan Wang et al [30]. Data citra sinar-x dada tersebut dapat diunduh di alamat https://github.com/muhammedtalo/COVID-19. Repositori data citra sinar-x COVID-19 dikembangkan oleh Cohen JP untuk mengumpulkan data dari berbagai sumber atau secara *crowdsourcing* dengan para peneliti dari berbagai negara. Repositori ini terus diperbarui sehingga total terdapat 1125 citra yang terdiri atas 125 citra pasien positif COVID-19, 500 citra normal, dan 500 citra Pneumonia. Dari 125 citra pasien yang positif COVID-19, terdapat 43 pasien yang berjenis kelamin perempuan dan 82 pasien yang berjenis kelamin laki-laki. Gambar 11 berturut-turut menunjukkan contoh citra pasien terpapar COVID-19, kondisinya normal, dan terpapar Pneumonia yang mana secara sepintas sulit dibedakan melalui mata telanjang.

Gambar 11. Dari kiri ke kanan: citra sinar-x dada pasien yang terpapar COVID-19, kondisi normal, dan terpapar pneumonia

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 86

Tabel 1. Perbandingan Akurasi dengan Model *Deep Learning* Lain

	Akurasi	# parameter
FullConv	86.93 %	10.951.059
ResNet50	90.8 %	23.514.179
EfficientNet	87.5 %	28.346.931
ShuffleNet	86.93 %	1.267.759

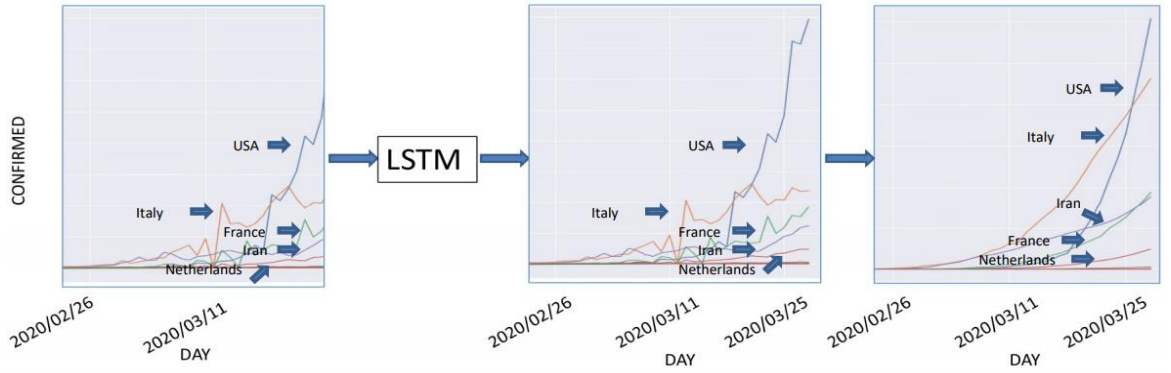


Gambar 12. Perbandingan Kecepatan Deteksi Diantara Arsitektur

Model CNN ShuffleNet yang kita ajukan mampu menghasilkan akurasi deteksi COVID-19 dengan akurasi sedikit dibawah dari model yang parameternya lebih banyak seperti EfficientNet dan ResNet50 (Tabel 1). Namun demikian, model yang diajukan mempunyai jumlah parameter yang jauh lebih sedikit 18.55 kali dari model lain sebelumnya yaitu EfficientNet dan 22.36 dari ResNet50. Selain itu, ShuffleNet menghabiskan memori GPU paling sedikit sebesar 0.646 GB (Gambar 12) serta waktu deteksi tercepat sebesar 0.0027 detik (Gambar 13), sehingga memungkinkan untuk diaplikasikan pada perangkat keras maupun aplikasi *mobile*.

K. Prediksi Laju COVID-19 dan Faktor-faktornya Melalui Data *Time-series* Multivariabel (domain Ekonomi, Politik, Alam, dan Sosial)

Dalam penelitian ini [31][32], sekumpulan data *time series* multivariabel dari 59 faktor dari 55 negara terdiri dari 6 perilaku, 21 COVID-19, 1 pemerintahan, 2 geografis, 3 morfologi, 2 ekonomi, 5 kesehatan, 2 fasilitas kesehatan, 1 pendidikan, dan 16 lingkungan. kategori. Setiap faktor ditangkap setiap hari, kecuali faktor lingkungan (UV) yang diwakili oleh observasi rata-rata harian. Fitur tersebut kemudian dilatih oleh varian DL yang disebut *Long Short Term Memory* (LSTM) [33] untuk memprediksi banyaknya orang yang terpapar COVID-19 hasil dari 59 faktor harian selama periode 174 hari (2020-03-22 hingga 2020-09-11) seperti yang nampak pada Gambar 14 pada berbagai negara. Untuk mencapai prediksi data spasial dan temporal tersebut, kami mengembangkan model *Convolution-LSTM* yang terdiri dari 3 lapisan tersembunyi dengan diikuti oleh unit linier dengan aktivasi Sigmoid. Seluruh jaringan terdiri dari unit LSTM yang berisi gerbang masukan, gerbang lupa, dan gerbang keluaran untuk menangkap korelasi spasial-temporal dan dinamika data deret waktu multivariabel.



Gambar 13. Kerangka Pelatihan dan Pengujian.
Masukan terdiri dari 67 hari dan keluarannya berupa prediksi terdiri dari 100 hari kasus harian.

Selama proses pembelajaran, pengoptimalan berbasis gradien melalui propagasi mundur digunakan. Dengan cara tersebut, peta atribusi (peta faktor) dibuat untuk memvisualisasikan fitur yang relevan dengan prediksi deret waktu spasial-temporal akhir. Secara khusus, metode yang disebut GradCAM [34] digunakan

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 87

untuk membuat peta saliensi yang berguna sebagai faktor penyebab dinamika pada sisi prediksi [35]. Grad-Cam diterapkan ke lapisan tersembunyi terakhir di mana aktivasi keluarannya dibobotkan dengan bobot penting yang terkait dengan prediksi deret waktu diikuti dengan aktivasi ReLU (*Rectified Linear Unit*).

4. Kesimpulan

Dengan analisa *Big Data* dan model pembelajaran *Deep Learning* sangat berpotensi untuk menggali pengetahuan secara komprehensif serta melakukan otomotisasi intelejensia pada komputer seperti halnya melakukan prediksi secara multidisplin dengan data yang besar. Berbeda dengan algoritma pembelajaran mesin dan rekayasa fitur yang lebih konvensional, *Deep Learning* memiliki keuntungan karena berpotensi memberikan solusi untuk menganalisis dan melakukan prediksi dengan set data dengan volume yang besar serta jumlah variabel dan tingkat kompleksitas yang tinggi. Lebih lanjut, masih banyak pekerjaan yang perlu dilakukan untuk bagaimana kita dapat mengadaptasi algoritma *Deep Learning* untuk masalah yang berkaitan dengan *Big Data*, termasuk akuisisi data berdimensi tinggi, analisis data *streaming* yang terus berubah dari sisi variasi dan kuantitasnya, skalabilitas model *Deep Learning* itu sendiri, serta komputasi yang efisien. Pekerjaan selanjutnya harus mulai memikirkan penanganan masalah secara komprehensif menggunakan data yang lebih besar, komplit, bekerjasama antar bidang dan pakar dengan menerapkan konsep *Big Data* dan model kecerdasan *Deep Learning* sehingga mampu untuk berkontribusi secara positif pada kemajuan di berbagai aspek bidang kehidupan terutama untuk Indonesia ke depannya.

5. Daftar Pustaka

[1] Syed, A., Gillela, K., & Venugopal, C. (2013). The future revolution on *Big Data*. *Future*, *2*(6), 2446-2451.

[2] Marjani, M., Nasaruddin, F., Gani, A., Karim, A., Hashem, I. A. T., Siddiqa, A., & Yaqoob, I. (2017). Big IoT data analytics: architecture, opportunities, and open research challenges. *IEEE Access*, *5*, 5247-5261.

[3] Sagiroglu, S., & Sinanc, D. (2013, May). *Big Data*: A review. In *2013 international conference on collaboration technologies and systems (CTS)* (pp. 42-47). IEEE.

[4] Erl, T., Khattak, W., & Buhler, P. (2016). *Big Data fundamentals: concepts, drivers & techniques*. Prentice Hall Press.

[5] Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013, January). *Big Data*: Issues and challenges moving forward. In *2013 46th Hawaii International Conference on System Sciences* (pp. 995-1004). IEEE.

[6] Mohammadi, M., Al-Fuqaha, A., Sorour, S., & Guizani, M. (2018). *Deep Learning* for IoT *Big Data* and streaming analytics: A survey. *IEEE Communications Surveys & Tutorials*, *20*(4), 2923-2960.

[7] Twitter Search oleh Marc Smith. https://blogs.lse.ac.uk/impactofsocialsciences/2015/07/10/social-media-research-tools-overview/

[8] Rodríguez, J. P., Fernández-Gracia, J., Thums, M., Hindell, M. A., Sequeira, A. M., Meekan, M. G., ... & Muelbert, M. (2017). *Big Data* analyses reveal patterns and drivers of the movements of southern elephant seals. *Scientific reports*, *7*(1), 1-10.

[9] Gaffney, D., & Matias, J. N. (2018). Caveat emptor, computational social science: Large-scale missing data in a widely-published Reddit corpus. *PloS one*, *13*(7), e0200162.

[10] LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep Learning*. *nature*, *521*(7553), 436-444.

[11] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition.

[12] Berniker, M., & Kording, K. P. (2015). Deep networks for motor control functions. *Frontiers in computational neuroscience*, *9*, 32.

[13] Bae, H. S., Lee, H. J., & Lee, S. G. (2016, June). Voice recognition based on adaptive MFCC and *Deep Learning*. In *2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA)* (pp. 1542-1546). IEEE.

[14] Kruger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J., ... & Wiskott, L. (2012). Deep hierarchies in the primate visual cortex: What can we learn for computer vision?. *IEEE transactions on pattern analysis and machine intelligence*, *35*(8), 1847-1871.

[15] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

[16] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.

[17] IBM. https://www.ibm.com/blogs/systems/deep-learning-performance-breakthrough/

[18] Soomro, K., Zamir, A. R., & Shah, M. (2012). UCF101: A dataset of 101 human actions classes from videos in the wild. *arXiv preprint arXiv:1212.0402*.

[19] Fukushima, K. (1988). Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural networks*, *1*(2), 119-130.

[20] Yudistira, N., & Kurita, T. (2017). Gated spatio and temporal convolutional neural network for activity recognition: towards gated multimodal *Deep Learning*. *EURASIP Journal on Image and Video Processing*, *2017*(1), 85.

[21] Ichinose, T. M., & Iwane, A. H. (2017). Cytological analyses by advanced electron microscopy. In *Cyanidioschyzon merolae* (pp. 129-151). Springer, Singapore.

[22] Yudistira, N., Kavitha, M., Itabashi, T., Iwane, A. H., & Kurita, T. (2020). prediction of Sequential organelles Localization under imbalance using A Balanced Deep U-net. *Scientific Reports*, *10*(1), 1-11.

[23] Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., ... & Yuan, M. L. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, *579*(7798), 265-269.

[24] Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., ... & Cheng, Z. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The lancet*, *395*(10223), 497-506.

[25] Wu, Z., & McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *Jama*, *323*(13), 1239-1242.

[26] M.L. Holshue, C. Debolt, et al. (2020). First case of 2019 novel coronavirus in the United States. N. Engl. J. Med. 328, p.929–936.

[27] Zhang, R., Li, Y., Zhang, A. L., Wang, Y., & Molina, M. J. (2020). Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proceedings of the National Academy of Sciences*.

[28] H. SHI, X. HAN, et al. (2020). Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. Lancet Infect. Dis. 24 (4), p.425–434.

[29] J.P. COHEN. (2020). COVID-19 Image Data Collection. https://github.com/ieee8023/COVID-chestxray-dataset.

[30] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers. 2017. Chestx-ray8: hospital scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.2097

[31] Novanto Yudistira. (2020) COVID-19 Growth Prediction using Multivariate Long Short-Term Memory. IAENG International Journal of Computer Science, vol. 47, no.4, pp829-837.

[32] Yudistira, N., Sumitro, S. B., Nahas, A., & Riama, N. F. (2020). UV light influences covid-19 activity through *Big Data*: tradeoffs between northern subtropical, tropical, and southern subtropical countries. *medRxiv*.

[33] Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.

[34] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision* (pp. 618-626).

[35] Yudistira, N., Sumitro, S. B., Nahas, A., & Riama, N. F. (2021). Learning where to look for COVID-19 growth: Multivariate analysis of COVID-19 cases over time using explainable convolution–LSTM. *Applied Soft Computing*, *109*, 107469.

DOI http://dx.doi.org/10.36448/expert.v11i2.2063
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 11 No. 2
December 27, 2021 – Hal. 89