

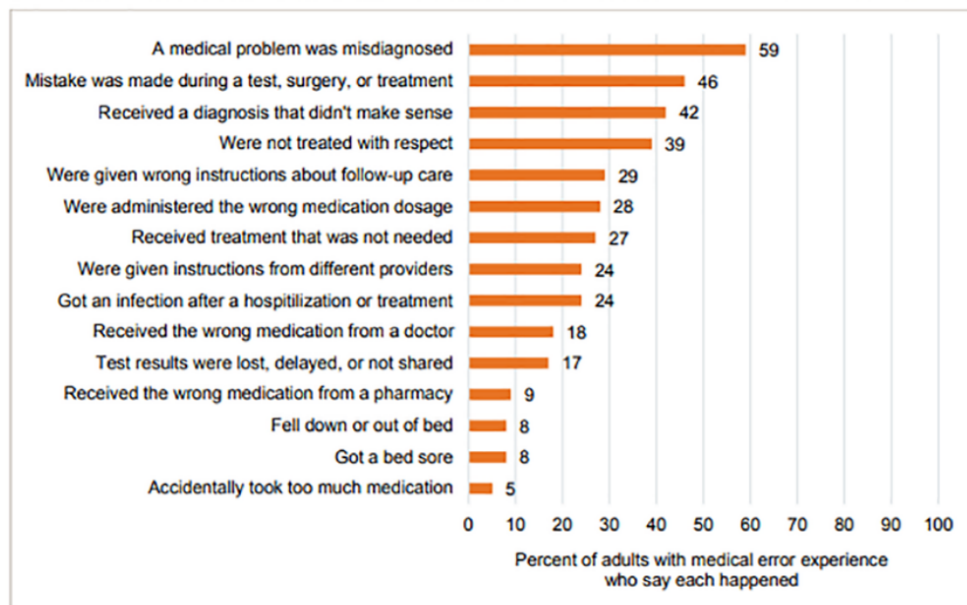
# BAB 1

## USULAN GAGASAN

### 1.1 Deskripsi Umum Masalah

Pada zaman digital ini, perkembangan teknologi telah membawa perubahan besar dalam berbagai aspek kehidupan, termasuk dalam dunia medis. Salah satu perubahan yang signifikan adalah kemampuan untuk menghasilkan dan menganalisis gambar medis dengan lebih cepat dan akurat. Namun, bersama dengan kemajuan ini, muncul pula tantangan baru, terutama dalam memproses dan menganalisis volume besar data gambar medis secara efisien.

Six in 10 adults with medical error experience say a medical problem was misdiagnosed and 4 in 10 say they weren't treated with respect.



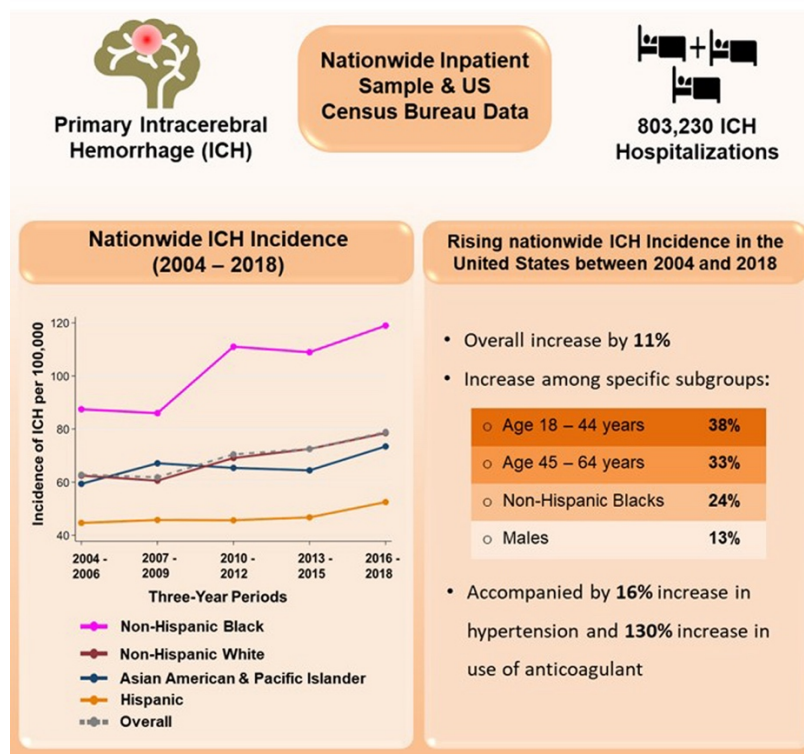
Question: Again, thinking about the most recent time a medical error was made in [your care/the care of someone close to you], for each of the following, please indicate whether or not it is the sort of medical error that occurred.

**Gambar 1.1 Data survei tentang pengalaman kesalahan medis yang dialami oleh orang dewasa**

Berdasarkan survei yang dilakukan Institute for Healthcare Improvement, seperti yang terlihat pada Gambar 1.1, enam dari sepuluh orang dewasa dengan pengalaman kesalahan medis melaporkan bahwa masalah medis mereka salah didiagnosis, dan empat dari sepuluh merasa tidak diperlakukan dengan hormat. Kesalahan medis lainnya termasuk menerima diagnosis yang tidak masuk akal (42%), kesalahan selama tes, operasi, atau perawatan (46%), serta menerima instruksi yang salah tentang perawatan lanjutan (29%) [1]. Kesalahan dalam diagnosis dan aspek terkait merupakan masalah signifikan dalam sistem perawatan kesehatan.

Pendeteksian yang cepat dan akurat pada penyakit sangat penting untuk mengurangi kesalahan ini, mencegah kesalahan perawatan, memastikan perawatan yang sesuai, dan meningkatkan kepercayaan pasien. Oleh karena itu, fokus pada peningkatan akurasi diagnosis dan percepatan pendeteksiannya adalah langkah kritis untuk meningkatkan keselamatan dan kualitas perawatan pasien.

Salah satu masalah kritis dalam praktik medis saat ini adalah deteksi dini dan penanganan yang cepat terhadap perdarahan intrakranial [2]. Perdarahan intrakranial yang selanjutnya disebut *Intracranial Hemorrhage (ICH)* adalah perdarahan yang terjadi di dalam ruang intrakranial, termasuk parenkim otak dan ruang meningeal di sekitarnya [3]. Meskipun telah ada teknologi pencitraan medis yang canggih, seperti *Computed Tomography (CT) scan* dan *Magnetic Resonance Imaging (MRI)*, akan tetapi, proses analisis gambar yang manual masih memakan waktu dan rentan terhadap kesalahan manusia [4]. Dalam situasi ICH, deteksi yang lambat atau kurang akurat dapat berakibat fatal bagi pasien.



**Gambar 1.2** Tren insiden Intracerebral Hemorrhage (ICH) di AS dari tahun 2004 hingga 2018 [5].

**Gambar 1.2** menyajikan data tentang kejadian ICH di Amerika Serikat berdasarkan sampel rawat inap nasional dan data Biro Sensus AS dari tahun 2004 hingga 2018, yang melibatkan 803.230 rawat inap ICH. Data menunjukkan bahwa insiden ICH meningkat sebesar 11%

selama periode tersebut, dengan peningkatan terbesar di kalangan usia 18-44 tahun (38%), usia 45-64 tahun (33%), *Non-Hispanic Black* (24%), dan pria (13%). Peningkatan ini disertai dengan kenaikan 16% dalam kasus hipertensi dan 130% dalam penggunaan antikoagulan. Kelompok non-Hispanik Hitam mengalami peningkatan insiden tertinggi dibandingkan kelompok demografis lainnya [5].

Akuisisi data CT menghasilkan sinogram, yang merupakan data mentah yang tidak dapat diinterpretasikan secara langsung oleh manusia. Sinogram harus direkonstruksi terlebih dahulu menjadi citra CT yang dapat dibaca oleh radiolog atau ahli melalui proses komputasi yang kompleks dan memakan waktu [6]. Interpretasi citra CT oleh radiolog atau dokter membutuhkan keahlian dan pengalaman yang tinggi. Proses ini rentan terhadap kesalahan manusia dan subjektivitas, serta dapat menyebabkan keterlambatan dalam diagnosis dan perawatan pasien [7]. Selain itu, terdapat kekurangan tenaga profesional medis yang terampil dalam membaca citra CT, terutama di daerah terpencil atau fasilitas kesehatan dengan sumber daya terbatas.

Untuk mengatasi tantangan-tantangan tersebut, pendekatan baru yang melibatkan *machine learning* telah dikembangkan. Penelitian terkini menunjukkan bahwa algoritma *deep learning* (yang merupakan subset dari *machine learning*) dapat dilatih untuk menginterpretasi citra CT secara langsung. Pendekatan ini berpotensi untuk mempercepat proses diagnosis, meningkatkan akurasi, dan mengurangi paparan radiasi dengan menggunakan sinogram yang diambil dengan dosis rendah.

Akuisisi data CT melibatkan paparan radiasi pengion yang dapat berbahaya bagi pasien jika dosis tidak dikelola dengan baik [8]. Upaya untuk mengurangi dosis radiasi sering kali berkompromi dengan kualitas citra yang dihasilkan. Kualitas citra yang buruk dapat mengurangi akurasi radiologi/ahli mendiagnosis [9].

Namun, pengembangan sistem pembelajaran mesin yang dapat menginterpretasi citra CT secara akurat dan andal masih memiliki tantangan tersendiri. Citra CT memiliki struktur data yang kompleks dan tidak intuitif bagi orang awam dan bahkan untuk dokter yang masih belum banyak berpengalaman menginterpretasikan citra CT, sehingga membutuhkan arsitektur jaringan saraf tiruan yang dirancang khusus untuk dapat mempelajari pola dan fitur yang relevan dari data tersebut sebagai alat bantu untuk mempermudah interpretasi oleh orang awam dan tenaga medis. Selain itu, dibutuhkan dataset yang besar dan terannotasi dengan baik untuk melatih model pembelajaran mesin, serta strategi pelatihan yang efektif untuk mengoptimalkan kinerja model.

Oleh karena itu, kami merasa bahwa pengembangan sistem otomatis untuk mendeteksi dan mengklasifikasi perdarahan intrakranial atau dalam istilah medis disebut *intracranial hemorrhage* (ICH) pada citra pemindaian otak sangat penting. Dengan menggunakan teknik-teknik pengolahan citra yang canggih, kami berharap dapat menciptakan solusi yang dapat membantu dokter dalam membuat keputusan yang lebih cepat dan lebih tepat, sehingga meningkatkan prognosis pasien dan mengurangi risiko komplikasi yang serius. Dalam konteks ini, kami tertarik untuk mengeksplorasi penggunaan *machine learning* (khususnya *deep learning*) sebagai dasar untuk pengembangan solusi deteksi dan klasifikasi yang efektif, terkhusus untuk studi kasus perdarahan intrakranial pada otak. Pendekatan ini berpotensi untuk memberikan solusi yang efisien, akurat, dan aman dalam diagnosis dan perawatan medis berbasis pencitraan CT, serta membuka peluang untuk pengembangan alat diagnosis berbasis CT yang lebih sederhana dan terjangkau di masa depan.

## 1.2 Analisa Masalah

Penelitian ini berfokus pada deteksi dan klasifikasi ICH menggunakan metode *Deep Learning*. Pada bagian ini, kami menganalisa masalah yang diangkat berdasarkan berbagai aspek, termasuk aspek teknis, ekonomi, lingkungan, dan kesehatan.

### 1.2.1 Aspek Teknis

Metode klasifikasi citra tradisional dan modern yang menggunakan input berbasis ruang gambar menghadapi beberapa tantangan teknis yang signifikan, terutama yang terkait dengan ketergantungan pada data besar dan kebutuhan komputasi yang tinggi.

- **Keterbatasan Rekonstruksi CT Konvensional**

- **Masalah:** Proses rekonstruksi CT dari sinogram membutuhkan waktu lama (hingga 15 menit) meskipun akuisisi data hanya memerlukan waktu sekitar 2 menit, selain itu algoritma rekonstruksi yang digunakan dapat menyebabkan *noise* pada gambar CT hasil rekonstruksi [10].
- **Implikasi:** Hal ini berisiko hilangnya informasi yang terbentuknya artefak selama proses rekonstruksi CT, seperti artefak *beam hardening* dan *scattering* sehingga proses interpretasi manual dapat menjadi subjektif dan rentan terhadap kesalahan manusia.

- **Keterbatasan Interpretasi Gambar CT secara manual**

- **Masalah:** Interpretasi gambar CT untuk deteksi dan klasifikasi perdarahan intrakranial (ICH) bergantung pada kemampuan radiologi/ahli dalam menganalisis gambar secara visual dan dibutuhkan waktu dan keahlian khusus untuk menginterpretasikan gambar CT dengan akurat [11].
- **Implikasi:** Proses Interpretasi langsung melalui gambar hasil CT oleh radiologi/ahli menghasilkan keputusan subjektif berdasarkan gambar hasil rekonstruksi, sehingga diperlukan bantuan kesimpulan secara matematis dari algoritma *deep learning* yang mempelajari pola berdasarkan pada data citra besar (*big data*) yang didapat sebagai landasan awal radiologi/ahli sebelum mendiagnosis lebih lanjut [7].
- **Variasi Geometris dalam Citra**
  - **Masalah:** Citra sering kali mengalami variasi geometris seperti translasi, rotasi, dan skala yang dapat mempengaruhi kinerja model klasifikasi [12].
  - **Implikasi:** Metode konvensional yang saat ini telah banyak dikembangkan untuk kasus lain memerlukan teknik *augmentasi* data ekstensif untuk mengatasi variasi ini, yang menambah kompleksitas dalam proses pelatihan dan dapat memperlambat proses pengembangan model.

### 1.2.2 Aspek Ekonomi

Proses diagnosis penyakit melalui citra CT yang hanya mengandalkan ahli radiologi terlatih membawa sejumlah implikasi ekonomi yang tidak dapat diabaikan.

- **Biaya Sumber Daya Manusia**
  - **Masalah:** Dibutuhkan radiologi/ahli yang terlatih dan berpengalaman untuk menginterpretasikan gambar CT dengan akurat, yang membutuhkan biaya kompensasi yang tinggi.
  - **Implikasi:** Kesalahan interpretasi dapat menyebabkan biaya perawatan dan pengobatan yang lebih tinggi bagi pasien [7].

### 1.2.3 Aspek Lingkungan

Penerapan teknologi komputasi intensif memiliki dampak lingkungan yang perlu dipertimbangkan.

- **Pemaparan Radiasi**
  - **Masalah:** Pemindaian CT menggunakan radiasi pengion (sinar-X) yang dapat membahayakan kesehatan jika terpapar dalam jumlah besar dan terlalu lama, sedangkan untuk menghasilkan citra CT yang optimal secara visual dibutuhkan radiasi yang besar.

- **Implikasi:** Semakin lama waktu pemindaian dan rekonstruksi, semakin besar paparan radiasi yang diterima pasien dan petugas medis, ini dapat membahayakan kesehatan dan lingkungan sekitar [8].

#### 1.2.4 Aspek Kesehatan

Meski tidak langsung, ada beberapa dampak kesehatan yang mungkin timbul terkait dengan penggunaan teknologi komputasi intensif.

- **Risiko Keterlambatan Diagnosis**

- **Masalah:** Waktu yang dibutuhkan untuk rekonstruksi CT dan interpretasi manual dapat menyebabkan keterlambatan dalam diagnosis ICH, yang merupakan kondisi mengancam jiwa.
- **Implikasi:** Keterlambatan diagnosis dapat menyebabkan penurunan angka harapan hidup pasien dan meningkatkan risiko komplikasi [2].

- **Risiko Kesalahan Diagnosis**

- **Masalah:** Interpretasi yang hanya mengandalkan keahlian radiologi/ahli dalam menganalisis gambar hasil CT rentan terhadap kesalahan subjektif dan *human error* [7].
- **Implikasi:** Kesalahan diagnosis dapat menyebabkan keterlambatan atau kesalahan dalam penanganan medis, yang dapat membahayakan kesehatan dan keselamatan pasien [4].

Analisis masalah yang diangkat dalam *capstone design* ini menunjukkan keterbatasan dalam deteksi dan klasifikasi ICH menggunakan metode konvensional berbasis rekonstruksi CT dan interpretasi manual, dan metode klasifikasi citra berbasis *machine learning* konvensional menghadapi beberapa tantangan serius dari berbagai aspek, termasuk teknis, ekonomi, lingkungan, dan kesehatan. Untuk mengatasi tantangan-tantangan ini memerlukan pendekatan yang lebih efisien dan berkelanjutan, baik dari sisi teknologi maupun implementasinya dalam kehidupan nyata.

### 1.3 Analisis Solusi yang Ada

Dalam mengkaji solusi-solusi yang sudah ada terkait deteksi dan klasifikasi perdarahan intrakranial pada otak, beberapa metode yang umum digunakan diidentifikasi dan dianalisis berdasarkan keunggulan, kekurangan, dan keterbatasannya. Pertama, metode *Convolutional Neural Network* (CNN) memiliki akurasi tinggi dan mampu menangani dataset besar untuk pelatihan efektif serta memerlukan sumber daya komputasi yang signifikan [12]. Selain itu, CNN rentan terhadap *overfitting*. *Overfitting* adalah situasi di mana sebuah model belajar

terlalu baik dari data pelatihan hingga menangkap noise atau detail-detail yang tidak relevan, yang menyebabkan kinerja yang buruk pada data baru yang tidak pernah dilihat sebelumnya (data uji) [13]. Kedua, metode *Support Vector Machine* (SVM) efektif dalam ruang dimensi tinggi dan bekerja baik dengan dataset kecil hingga menengah. Namun, kesulitan dalam menangani dataset besar dan pemilihan kernel yang tepat menjadi tantangan utama, selain proses optimasi parameternya yang kompleks [14].

Metode lain seperti *Random Forest* menawarkan keunggulan dalam mengatasi *outlier* dan menghasilkan hasil yang baik melalui kombinasi model berbeda. Meski demikian, metode ini bisa lambat dan membutuhkan banyak memori jika jumlah *decision trees* yang digunakan sangat banyak, serta cenderung *overfitting* pada dataset kecil. Keterbatasannya terletak pada kinerja yang kurang optimal saat bekerja dengan data yang memiliki dimensi yang tinggi tanpa melakukan reduksi dimensi terlebih dahulu [15]. Sementara itu, *K-Nearest Neighbors* (KNN) memiliki implementasi yang sederhana dan tidak memerlukan asumsi distribusi data. Namun, kinerja KNN menurun pada dataset besar karena memerlukan banyak memori dan komputasi, serta sensitif terhadap skala data dan *outlier*.

Metode terbaru seperti *Radon Cumulative Distribution Transform* dan *Nearest Subspace* (RCDT-NS) menunjukkan pendekatan inovatif dengan hasil yang kompetitif [12]. Keunggulan metode ini termasuk kemampuan menangani variasi dalam bentuk dan ukuran perdarahan dengan baik. Namun, metode ini masih dalam tahap penelitian dan pengembangan, sehingga referensi yang mendukung belum banyak, dan memerlukan pemahaman mendalam tentang transformasi Radon dan distribusi kumulatif. Keterbatasan lainnya adalah ketersediaan data latih yang terbatas dapat mempengaruhi performa model, serta adaptasi metode ini ke dalam sistem klinis memerlukan validasi yang ekstensif.

Dari solusi-solusi di atas, beberapa *pain points* yang diidentifikasi antara lain kebutuhan data latih yang besar, kompleksitas dan sumber daya komputasi yang tinggi, risiko *overfitting* dan sensitivitas terhadap *hyperparameter*, serta tantangan implementasi di dunia nyata. Berdasarkan *pain points* tersebut, solusi yang diusulkan harus fokus pada efisiensi data, optimasi sumber daya komputasi, prediksi yang akurat, peningkatan *robustness* modul terhadap variasi data dan *outlier*, serta validasi klinis yang mendalam.

## 1.4 Kesimpulan dan Ringkasan CD-1

Masalah deteksi dan klasifikasi perdarahan intrakranial pada otak merupakan isu yang sangat penting dalam dunia medis, terutama dalam konteks diagnosis cepat dan akurat yang krusial untuk keselamatan pasien. Urgensi pada masalah ini diperkuat oleh fakta bahwa perdarahan intrakranial merupakan kondisi yang mengancam jiwa dan memerlukan penanganan segera [16]. Dalam situasi darurat, setiap menit berharga, sehingga keterlambatan dalam diagnosis dapat berakibat fatal.

Kompleksitas masalah ini terletak pada berbagai tantangan teknis yang dihadapi dalam metode konvensional. Proses rekonstruksi citra CT dari sinogram membutuhkan waktu lama dan berisiko menghasilkan artefak yang dapat mengganggu interpretasi citra jika dosis radiasi yang diberikan minim [10]. Selain itu, interpretasi manual oleh radiolog memerlukan keahlian khusus dan rentan terhadap kesalahan manusia, yang dapat mengakibatkan keputusan diagnosis yang tidak konsisten dan subjektif [11]. Tantangan ini semakin diperparah oleh keterbatasan jumlah data pelatihan yang diperlukan oleh metode *deep learning*, serta kebutuhan komputasi yang tinggi yang dapat menjadi penghalang di lingkungan dengan sumber daya terbatas.

Solusi-solusi yang sudah ada, seperti metode CNN, SVM, *Random Forest*, RCDT-NS dan KNN, memiliki keunggulan masing-masing, tetapi juga dihadapkan pada sejumlah keterbatasan atau pain points. CNN, misalnya, membutuhkan jumlah data pelatihan yang besar dan komputasi yang tinggi, sementara SVM kesulitan menangani dataset besar dan memerlukan optimasi kernel yang tepat [12], [14]. *Random Forest* efektif mengatasi *outlier*, tetapi rentan terhadap *overfitting* pada dataset kecil. Lalu, RCDT-NS yang performa modelnya di pengaruhi oleh ketersediaan data latih dan KNN yang memiliki masalah performa pada dataset besar serta sensitif terhadap skala data.