

MEDACT: PERANCANGAN DAN ANALISIS TEORITIS AKTIVASI BARU UNTUK DEEP NEURAL NETWORK PADA CITRA MEDIS

Muhammad Syarif Baital.
 STMIK Catur Sakti Kendari
e-mail: syarif.baital@gmail.com

Fungsi aktivasi merupakan komponen fundamental dalam deep neural network karena menentukan kemampuan model dalam mempelajari hubungan nonlinier dari data. Dalam konteks citra medis, fungsi aktivasi tidak hanya dituntut stabil secara numerik tetapi juga mampu mengekstraksi fitur morfologi halus dengan sensitivitas tinggi. Artikel ini mengusulkan sebuah fungsi aktivasi baru bernama MedAct (Medical Activation Function) yang dirancang secara teoritis sebagai fungsi hibrida antara komponen linear, sigmoid, dan hiperbolik. MedAct dirumuskan sebagai $MedAct(x) = x \cdot \tanh(\alpha \cdot \sigma(x))$, dengan parameter alpha (α) yang dapat ditetapkan secara tetap (fixed) maupun dipelajari (learnable). Penelitian ini membahas prinsip perancangan, formulasi matematis, sifat diferensiabilitas, stabilitas gradien, analisis perilaku asimtotik, serta relevansinya terhadap karakteristik citra medis. Kajian ini bersifat konseptual dan dimaksudkan sebagai dasar teoritis bagi penelitian eksperimental lanjutan pada arsitektur deep learning seperti ResNet..

Kata Kunci — fungsi aktivasi, MedAct, deep learning, citra medis, analisis matematis.

I. PENDAHULUAN

Perkembangan *deep learning* dalam satu dekade terakhir telah membawa transformasi signifikan dalam bidang pengolahan citra medis. Berbagai tugas seperti klasifikasi kanker, segmentasi organ, deteksi malformasi anatomi, serta prediksi penyakit berbasis citra menunjukkan peningkatan akurasi yang substansial dibandingkan metode konvensional. Hal ini didorong oleh kemajuan arsitektur *deep neural network*, khususnya *Convolutional Neural Network* (CNN) dan *Residual Network* (ResNet), yang mampu mengekstraksi fitur kompleks dari data berdimensi tinggi secara otomatis [1], [2], [3], [4], [5], [6].

Model *deep learning* saat ini tidak hanya digunakan sebagai alat bantu diagnosis, tetapi juga mulai diintegrasikan dalam sistem pendukung keputusan klinis. Dalam beberapa studi terbaru, performa *deep learning* bahkan mendekati atau melampaui tingkat akurasi pakar medis pada tugas-tugas tertentu seperti deteksi tumor dan

klasifikasi penyakit berbasis citra radiologi dan histopatologi [7], [8], [9], [10], [11], [12], [13]. Fenomena ini menunjukkan bahwa *deep learning* memiliki potensi besar dalam meningkatkan efisiensi dan objektivitas proses diagnosis medis.

Keberhasilan model *deep learning* tidak hanya ditentukan oleh desain arsitektur jaringan, tetapi juga oleh komponen fundamental di dalamnya, salah satunya adalah fungsi aktivasi. Fungsi aktivasi berperan sebagai mekanisme nonlinier yang mengubah sinyal linear hasil operasi konvolusi menjadi representasi yang lebih kompleks. Tanpa fungsi aktivasi yang tepat, jaringan saraf hanya akan berperilaku sebagai transformasi linear sehingga tidak mampu memodelkan hubungan kompleks pada data citra medis [14], [15], [16].

Pemilihan fungsi aktivasi yang kurang sesuai dapat menimbulkan berbagai permasalahan selama proses pelatihan jaringan, seperti *vanishing gradient*, *exploding gradient*, dan fenomena *dead neuron*. Masalah-masalah ini berpotensi menghambat konvergensi model dan menurunkan kemampuan jaringan dalam mengekstraksi fitur penting dari citra medis. Oleh karena itu, fungsi aktivasi menjadi komponen krusial yang secara langsung mempengaruhi stabilitas numerik dan performa akhir model *deep learning* [17].

Dalam domain citra medis, tantangan yang dihadapi jauh lebih kompleks dibandingkan citra alami. Citra medis sering memiliki kontras rendah antar jaringan, *noise* yang berasal dari proses akuisisi, serta struktur anatomi dengan perbedaan intensitas yang sangat halus. Selain itu, distribusi data pada *dataset* medis sering kali tidak seimbang karena keterbatasan jumlah sampel penyakit tertentu. Kondisi ini menuntut pendekatan pemodelan yang lebih sensitif terhadap variasi kecil pada intensitas piksel dan struktur morfologi [4], [5], [6], [18], [19], [20].

Sebagian besar fungsi aktivasi populer seperti ReLU, ELU, Swish, dan Mish dirancang untuk keperluan umum pada data citra alami. ReLU, misalnya, memotong seluruh nilai negatif sehingga berpotensi menghilangkan informasi penting yang merepresentasikan perbedaan jaringan halus pada citra medis. Sementara itu, fungsi aktivasi *smooth* seperti Swish dan Mish menawarkan keunggulan dalam hal

kelicinan gradien, namun tetap tidak secara khusus mempertimbangkan karakteristik domain medis [21], [22], [23].

Beberapa penelitian terbaru menunjukkan bahwa fungsi aktivasi yang lebih halus (*smooth activation functions*) cenderung memberikan hasil yang lebih stabil pada aplikasi citra medis dibandingkan fungsi aktivasi berbasis ambang keras (*hard threshold*) seperti ReLU. Aktivasi yang halus memungkinkan jaringan mempertahankan informasi pada domain negatif dan merespons perubahan kecil pada intensitas piksel secara lebih adaptif. Hal ini sangat penting pada kasus-kasus klinis di mana perbedaan antar kelas sangat tipis dan sulit dibedakan secara visual [24], [25], [26], [27].

Meskipun demikian, sebagian besar penelitian masih berfokus pada pemilihan fungsi aktivasi yang telah ada, bukan pada perancangan fungsi aktivasi baru yang secara khusus dioptimalkan untuk domain medis. Kesenjangan penelitian ini membuka peluang untuk mengembangkan fungsi aktivasi yang dirancang berdasarkan karakteristik biologis dan morfologis citra medis. Pendekatan ini diharapkan dapat meningkatkan kualitas representasi fitur serta stabilitas pelatihan *deep neural network* [28], [29], [30], [31].

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan sebuah fungsi aktivasi baru bernama MedAct (*Medical Activation Function*) yang dirancang secara teoritis dengan menggabungkan sifat linearitas terkontrol, modulasi nonlinier, serta stabilitas gradien. MedAct memanfaatkan kombinasi fungsi sigmoid dan hiperbolik sebagai mekanisme gating nonlinier yang mengatur kontribusi input terhadap *output neuron*. Desain ini diharapkan mampu mempertahankan sensitivitas terhadap variasi intensitas kecil sekaligus menjaga stabilitas numerik selama proses pelatihan jaringan.

Artikel ini berfokus pada pengembangan konseptual dan analisis matematis MedAct sebagai dasar ilmiah sebelum dilakukan validasi empiris pada penelitian berikutnya. Kajian ini mencakup formulasi matematis, sifat diferensiabilitas, analisis perilaku asimtotik, serta stabilitas gradien MedAct. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi teoretis dalam pengembangan fungsi aktivasi baru yang lebih sesuai untuk aplikasi *deep learning* pada citra medis.

II. TINJAUAN PUSTAKA

A. Fungsi Aktivasi dalam Neural Network

Fungsi aktivasi diperkenalkan untuk mengatasi keterbatasan model linear dalam *neural network*. Tanpa fungsi aktivasi nonlinier, *neural network* hanya mampu merepresentasikan transformasi linear dari *input* sehingga tidak dapat memodelkan hubungan kompleks pada data. Dalam *deep learning* modern, fungsi aktivasi berperan penting dalam menentukan kapasitas representasi dan stabilitas pelatihan jaringan [14], [15].

Rectified Linear Unit (ReLU) menjadi fungsi aktivasi standar karena kesederhanaan formulasi dan efisiensi komputasinya. ReLU mempercepat konvergensi pelatihan

dan mengurangi permasalahan *vanishing gradient* pada *deep neural network*. Namun, ReLU memiliki kelemahan utama berupa fenomena *dead neuron*, yaitu kondisi ketika *neuron* tidak lagi aktif karena gradien bernilai nol pada domain negatif secara permanen [17], [32], [33].

Untuk mengatasi kelemahan tersebut, beberapa fungsi aktivasi alternatif telah dikembangkan. *Exponential Linear Unit* (ELU) dan *Scaled Exponential Linear Unit* (SELU) memperkenalkan domain negatif yang halus sehingga *neuron* tetap aktif selama proses pelatihan. Fungsi-fungsi ini juga bertujuan menjaga stabilitas distribusi aktivasi di dalam jaringan, khususnya pada arsitektur yang lebih dalam [34], [35], [36], [37], [38], [39].

Selain itu, fungsi aktivasi modern seperti Swish dan Mish diperkenalkan dengan sifat non-monotonik dan *smooth* untuk meningkatkan stabilitas gradien. Swish dirumuskan sebagai fungsi berbasis sigmoid yang memodulasi *input* secara adaptif, sedangkan Mish mengombinasikan fungsi hiperbolik dan logaritmik untuk menghasilkan kurva aktivasi yang lebih halus. Beberapa penelitian menunjukkan bahwa fungsi aktivasi *smooth* mampu memberikan performa yang lebih baik pada berbagai tugas klasifikasi dan segmentasi dibanding ReLU [21], [22], [23].

Perkembangan fungsi aktivasi menunjukkan bahwa desain fungsi nonlinier memiliki pengaruh signifikan terhadap performa *neural network*. Namun, sebagian besar fungsi aktivasi yang ada dirancang secara umum tanpa mempertimbangkan karakteristik khusus dari domain aplikasi tertentu. Hal ini membuka peluang untuk mengembangkan fungsi aktivasi baru yang dioptimalkan berdasarkan kebutuhan spesifik suatu domain, termasuk domain citra medis yang memiliki tantangan tersendiri dalam representasi fitur [17], [28], [29], [30], [31].

B. Fungsi Aktivasi pada Aplikasi Citra Medis

Dalam aplikasi citra medis, peran fungsi aktivasi menjadi lebih kritis dibandingkan citra alami karena kesalahan kecil dalam representasi fitur dapat berdampak besar pada keputusan klinis. Model *deep learning* harus mampu membedakan variasi intensitas yang sangat halus antar jaringan biologis, seperti perbedaan antara jaringan sehat dan jaringan patologis. Oleh karena itu, sensitivitas fungsi aktivasi terhadap perubahan kecil pada *input* menjadi faktor penting dalam performa sistem diagnosis berbasis citra [1], [2], [31].

Beberapa studi menunjukkan bahwa fungsi aktivasi yang bersifat *smooth* cenderung menghasilkan representasi fitur yang lebih stabil pada *dataset* medis dibandingkan fungsi aktivasi berbasis *hard threshold* seperti ReLU. Aktivasi *smooth* memungkinkan gradien mengalir secara kontinu sehingga jaringan lebih mudah mempelajari pola-pola kompleks yang terdapat pada citra medis dengan kontras rendah [11], [13], [24].

Selain itu, citra medis sering mengandung *noise* yang berasal dari proses akuisisi seperti pencitraan radiologi, MRI, atau histopatologi digital. Fungsi aktivasi yang terlalu agresif dalam memotong nilai negatif berpotensi menghilangkan informasi penting yang berkaitan dengan struktur morfologi halus. Oleh karena itu, pendekatan aktivasi yang mempertahankan informasi pada domain

negatif secara halus menjadi lebih sesuai untuk aplikasi medis [26], [27], [40], [41].

Meskipun demikian, sebagian besar penelitian masih menggunakan fungsi aktivasi generik seperti ReLU, Leaky ReLU, atau Swish tanpa melakukan modifikasi khusus untuk domain medis. Hal ini menunjukkan bahwa pemilihan fungsi aktivasi sering dianggap sebagai komponen sekunder dibandingkan desain arsitektur jaringan. Padahal, beberapa penelitian terbaru menegaskan bahwa fungsi aktivasi memiliki kontribusi signifikan terhadap stabilitas pelatihan dan kualitas fitur yang dihasilkan [17], [22].

Dengan meningkatnya penggunaan deep learning dalam sistem klinis, kebutuhan akan fungsi aktivasi yang lebih sesuai dengan karakteristik citra medis menjadi semakin mendesak. Fungsi aktivasi yang dirancang khusus diharapkan mampu meningkatkan sensitivitas terhadap variasi morfologi kecil sekaligus menjaga stabilitas numerik selama pelatihan. Hal ini menjadi dasar penting bagi pengembangan fungsi aktivasi berbasis domain medis seperti MedAct yang diusulkan dalam penelitian ini [26], [27], [28], [30], [31].

C. Kesenjangan Penelitian

Meskipun berbagai fungsi aktivasi telah dikembangkan, sebagian besar penelitian masih berfokus pada pemilihan fungsi aktivasi yang telah ada daripada merancang fungsi aktivasi baru yang disesuaikan dengan domain tertentu. Dalam konteks citra medis, pendekatan ini berpotensi mengabaikan karakteristik khusus data seperti kontras rendah dan variasi intensitas yang halus [1], [2], [24], [42].

Kesenjangan pertama terletak pada kurangnya fungsi aktivasi yang dirancang berbasis domain medis. Sebagian besar fungsi aktivasi modern, termasuk Swish dan Mish, dikembangkan melalui pendekatan umum tanpa mempertimbangkan sifat biologis atau morfologis citra medis. Padahal, desain fungsi aktivasi berbasis domain dapat memberikan keuntungan dalam representasi fitur dan stabilitas pelatihan [26], [28], [31], [32].

Kesenjangan kedua adalah minimnya kajian teoritis mendalam terhadap fungsi aktivasi baru. Banyak penelitian memperkenalkan fungsi aktivasi berdasarkan eksperimen empiris tanpa analisis matematis yang komprehensif terkait sifat diferensiabilitas, stabilitas gradien, dan perilaku asimtotik. Padahal, analisis teoritis sangat penting untuk memahami mekanisme kerja fungsi aktivasi dalam *deep neural network* [15], [17].

Kesenjangan ketiga berkaitan dengan fokus penelitian yang lebih besar pada pengembangan arsitektur jaringan dibandingkan pada komponen dasar seperti fungsi aktivasi. Penelitian sering menitikberatkan pada desain CNN atau ResNet yang lebih dalam dan kompleks, sementara fungsi aktivasi dianggap sebagai komponen tetap yang tidak perlu dikaji ulang. Pendekatan ini membatasi potensi peningkatan performa yang dapat dicapai melalui inovasi pada tingkat fungsi aktivasi [9], [10], [11], [13], [22], [23].

Berdasarkan kesenjangan tersebut, MedAct diusulkan sebagai fungsi aktivasi baru yang dirancang melalui pendekatan teoritis yang sistematis dan terinspirasi oleh kebutuhan domain citra medis. MedAct tidak hanya

memperkenalkan formulasi matematis baru, tetapi juga menyediakan landasan analisis matematis yang kuat sebelum dilakukan validasi empiris. Dengan demikian, penelitian ini diharapkan dapat mengisi celah antara kebutuhan praktis aplikasi medis dan pengembangan teori fungsi aktivasi dalam *deep learning* [26], [27], [28], [29], [30], [31].

III. METODE PENELITIAN

A. Prinsip Dasar Perancangan

Perancangan fungsi aktivasi MedAct didasarkan pada kebutuhan khusus domain citra medis yang menuntut kestabilan numerik dan sensitivitas tinggi terhadap variasi intensitas kecil. Studi terbaru menunjukkan bahwa fungsi aktivasi memiliki pengaruh signifikan terhadap kualitas fitur yang dipelajari *deep neural network* pada data medis yang bersifat heterogen dan berisik [17], [24], [25].

Prinsip pertama dalam desain MedAct adalah stabilitas numerik. Fungsi aktivasi harus memiliki gradien yang tidak mengalami diskontinuitas atau saturasi ekstrem yang dapat menyebabkan *vanishing gradient* atau *exploding gradient*. Hal ini penting karena model *deep learning* untuk citra medis sering kali memiliki kedalaman jaringan yang besar untuk menangkap struktur anatomi kompleks [32], [33].

Prinsip kedua adalah nonlinearitas adaptif. MedAct dirancang agar mampu memberikan respons nonlinier yang fleksibel terhadap perubahan kecil pada nilai input. Pada citra medis, perbedaan intensitas antar jaringan sering kali sangat halus, sehingga fungsi aktivasi harus mampu mempertahankan informasi tersebut tanpa memotongnya secara keras sebagaimana terjadi pada ReLU [26], [27], [43], [44].

Prinsip ketiga adalah kesesuaian dengan karakteristik citra medis. MedAct tidak hanya mempertimbangkan aspek matematis, tetapi juga aspek semantik data medis, yaitu pentingnya kontinuitas, kehalusan, dan sensitivitas terhadap detail morfologi. Pendekatan ini sejalan dengan tren terkini yang mengusulkan fungsi aktivasi berbasis domain (*domain-specific activation functions*) untuk meningkatkan relevansi klinis model [28], [29], [30], [31].

Dengan menggabungkan ketiga prinsip tersebut, MedAct diposisikan sebagai fungsi aktivasi yang dirancang secara konseptual untuk kebutuhan citra medis, bukan sekadar adaptasi dari fungsi aktivasi umum. Pendekatan ini menekankan bahwa desain fungsi aktivasi harus dipandang sebagai komponen metodologis utama, bukan hanya pilihan teknis minor dalam arsitektur jaringan.

B. Formula Matematis

Secara matematis, MedAct dirumuskan sebagai

$$\text{MedAct}(x) = x \cdot \tanh(\alpha \cdot \sigma(x))$$

dengan $\sigma(x)$ merupakan fungsi sigmoid dan α adalah parameter pengendali nonlinieritas. Formulasi ini memungkinkan modulasi nonlinier terhadap sinyal linear x melalui fungsi yang halus dan terbatas (*bounded*).

Komponen x mempertahankan sifat linearitas yang

memungkinkan propagasi informasi tanpa distorsi berlebihan. Hal ini penting untuk menjaga sensitivitas terhadap nilai intensitas kecil yang sering muncul pada citra medis seperti radiografi, CT-scan, atau MRI [12], [13].

Fungsi sigmoid $\sigma(x)$ berperan sebagai mekanisme normalisasi yang memetakan nilai input ke rentang [0,1]. Dengan demikian, nilai input yang ekstrem tidak langsung menghasilkan respons nonlinier yang berlebihan, sehingga meningkatkan kestabilan pelatihan jaringan [23].

Fungsi *tanh* memberikan nonlinieritas yang *smooth* dan simetris di sekitar nol, serta memiliki gradien yang kontinu di seluruh domain *real*. Kehalusan ini memudahkan proses optimasi berbasis gradien dan mengurangi risiko *dead neuron* yang sering terjadi pada ReLU [26], [27], [32], [33].

Dengan mengombinasikan ketiga komponen tersebut, MedAct menghasilkan fungsi aktivasi yang adaptif, stabil, dan terdiferensialkan secara penuh. Struktur matematis ini diharapkan mampu meningkatkan kualitas representasi fitur pada data citra medis yang memiliki variasi intensitas halus dan distribusi kompleks.

C. Interpretasi sebagai Fungsi Gating

MedAct dapat ditulis ulang sebagai

$$\text{MedAct}(x) = x \cdot g(x)$$

dengan

$$g(x) = \tanh(\alpha \cdot \sigma(x)).$$

Dalam formulasi ini, fungsi $g(x)$ bertindak sebagai mekanisme *gating* adaptif yang mengontrol kontribusi sinyal input terhadap *output neuron*.

Konsep *gating* telah banyak digunakan dalam *neural network* modern, seperti pada LSTM, GRU, dan Swish activation, untuk mengatur aliran informasi secara selektif [22]. MedAct mengadopsi konsep serupa tetapi dirancang dengan struktur yang lebih sederhana dan kontinu.

Gating pada MedAct bersifat nonlinier dan halus karena dikendalikan oleh *sigmoid* dan *tanh*. Hal ini memungkinkan *neuron* untuk menyesuaikan tingkat aktivasi berdasarkan konteks nilai input, bukan hanya berdasarkan tanda positif atau negatif seperti pada ReLU.

Dalam konteks citra medis, mekanisme *gating* ini sangat relevan karena memungkinkan model mempertahankan fitur dengan intensitas rendah namun signifikan secara klinis, seperti batas jaringan atau anomali kecil pada struktur organ [24], [28], [31].

Dengan interpretasi *gating* ini, MedAct dapat dipahami tidak hanya sebagai fungsi aktivasi matematis, tetapi juga sebagai mekanisme seleksi fitur adaptif yang berpotensi meningkatkan sensitivitas model terhadap detail morfologi penting dalam citra medis.

D. Varian MedAct

MedAct memiliki dua varian utama, yaitu *Fixed- α* dan *Learnable- α* . Kedua varian ini dikembangkan untuk mengeksplorasi *trade-off* antara stabilitas pelatihan dan fleksibilitas adaptasi fungsi aktivasi terhadap data.

Pada varian *Fixed- α* , parameter α ditetapkan sebagai konstanta sepanjang proses pelatihan. Pendekatan ini

bertujuan menjaga kestabilan nonlinieritas serta memudahkan reproduktibilitas hasil penelitian, sebagaimana pendekatan parameter tetap pada fungsi aktivasi ELU dan SELU [34], [35], [36], [37], [38], [39].

Varian *Learnable- α* memungkinkan α dipelajari secara otomatis melalui *backpropagation* bersama bobot jaringan. Pendekatan ini memberikan fleksibilitas lebih tinggi karena fungsi aktivasi dapat menyesuaikan tingkat nonlinieritas sesuai karakteristik data yang dipelajari [15], [16], [17].

Namun, penggunaan *Learnable- α* memerlukan kontrol tambahan agar nilai α tidak tumbuh terlalu besar dan menyebabkan instabilitas gradien. Oleh karena itu, diperlukan strategi inisialisasi yang tepat atau regularisasi parameter agar proses pelatihan tetap konvergen [9], [32].

Perbandingan antara *Fixed- α* dan *Learnable- α* membuka peluang analisis lebih lanjut mengenai pengaruh parameterisasi fungsi aktivasi terhadap performa model, khususnya pada domain citra medis yang memiliki karakteristik distribusi data yang unik.

Pada tahap awal perancangan, MedAct dirumuskan sebagai fungsi hibrida kompleks yang mengombinasikan beberapa fungsi aktivasi populer, yaitu ReLU, Swish, dan Mish. Formulasi awal tersebut bertujuan mengeksplorasi kemungkinan peningkatan performa dengan menggabungkan karakteristik beberapa fungsi aktivasi yang telah terbukti efektif.

Namun, formulasi awal tersebut memiliki kompleksitas matematis yang tinggi dan sulit dianalisis secara teoritis. Selain itu, kombinasi beberapa fungsi aktivasi sekaligus berpotensi menimbulkan redundansi nonlinieritas serta meningkatkan risiko instabilitas gradien pada jaringan yang sangat dalam [32].

Oleh karena itu, formulasi MedAct kemudian disederhanakan menjadi $\text{MedAct}(x) = x \cdot \tanh(\alpha \cdot \sigma(x))$ yang mempertahankan prinsip *gating* nonlinier namun dengan struktur matematis yang lebih elegan dan mudah dianalisis secara formal.

Perubahan formulasi ini dilakukan untuk menekankan pendekatan teoritis yang sistematis, sehingga sifat-sifat matematis MedAct seperti kontinuitas, diferensiabilitas, dan stabilitas gradien dapat dianalisis secara lebih mendalam. Pendekatan ini sejalan dengan tren penelitian terbaru yang menekankan pentingnya justifikasi matematis dalam desain fungsi aktivasi baru [17], [28], [31].

Dengan demikian, perbedaan formulasi bukan merupakan inkonsistensi, melainkan hasil proses pematangan konsep dari eksplorasi empiris menuju formulasi teoritis yang lebih kuat dan dapat dijadikan dasar ilmiah untuk penelitian eksperimental selanjutnya.

IV. HASIL DAN PEMBAHASAN

A. Kontinuitas dan Diferensiabilitas

MedAct didefinisikan sebagai komposisi dari fungsi linear x , fungsi sigmoid $\sigma(x)$, dan fungsi hiperbolik $\tanh(x)$. Ketiga fungsi tersebut diketahui bersifat kontinu dan terdiferensialkan pada seluruh domain bilangan *real*.

Berdasarkan teorema komposisi fungsi kontinu dan terdiferensialkan, maka MedAct juga bersifat kontinu dan terdiferensialkan di seluruh domain *real* tanpa adanya titik diskontinuitas.

Sifat diferensiabilitas penuh ini sangat penting dalam konteks optimasi berbasis gradien, khususnya algoritma *backpropagation*. Berbeda dengan ReLU yang tidak terdiferensialkan di titik nol, MedAct tidak memiliki titik sudut sehingga gradien dapat dihitung secara konsisten pada setiap nilai *input* [32], [33].

Kontinuitas MedAct juga memastikan bahwa perubahan kecil pada *input* menghasilkan perubahan kecil pada *output*. Hal ini relevan untuk citra medis yang sering mengandung variasi intensitas sangat halus, seperti perbedaan jaringan lunak pada MRI atau CT-scan. Fungsi aktivasi yang tidak kontinu berpotensi memperbesar fluktuasi kecil menjadi perubahan besar pada representasi fitur.

Beberapa penelitian terbaru menekankan bahwa fungsi aktivasi *smooth* (kontinu dan terdiferensialkan penuh) memberikan stabilitas optimasi yang lebih baik dibanding fungsi aktivasi *piecewise* seperti ReLU, terutama pada jaringan yang sangat dalam [7], [12], [26], [27].

Dengan demikian, sifat kontinuitas dan diferensiabilitas MedAct menjadi landasan matematis utama yang mendukung klaim stabilitas gradien dan kompatibilitasnya dengan algoritma optimasi modern dalam *deep learning*, khususnya pada aplikasi citra medis yang sensitif terhadap noise dan variasi kecil.

B. Turunan Lengkap MedAct

Turunan pertama MedAct terhadap x adalah

$$\text{MedAct}'(x) = \tanh(\alpha \cdot \sigma(x)) + x(1 - \tanh^2(\alpha \cdot \sigma(x)))\alpha \cdot \sigma(x)(1 - \sigma(x)).$$

Eksresi ini menunjukkan bahwa gradien MedAct merupakan kombinasi dari dua komponen, turunan langsung dari faktor *tanh* dan kontribusi dari faktor linear x .

Komponen pertama, $\tanh(\alpha \cdot \sigma(x))$, menjamin bahwa gradien selalu memiliki kontribusi non-nol pada hampir seluruh domain, kecuali pada kondisi ekstrem ketika fungsi *tanh* mencapai saturasi. Hal ini mengurangi risiko *dead neuron* yang umum terjadi pada ReLU ketika *input* berada di domain negatif [32], [33].

Komponen kedua melibatkan faktor $\sigma(x)(1 - \sigma(x))$, yang merupakan turunan sigmoid dan memiliki nilai maksimum 0,25. Faktor ini memastikan bahwa kontribusi tambahan terhadap gradien tidak tumbuh secara tak terkendali, sehingga gradien tetap berada dalam batas yang stabil [22].

Struktur turunan MedAct menyerupai fungsi aktivasi berbasis *gating* seperti Swish, di mana gradien dikontrol oleh fungsi sigmoid. Namun, MedAct menambahkan lapisan nonlinier *tanh* yang lebih halus dan *bounded*, sehingga memberikan kontrol tambahan terhadap dinamika gradien [26], [27].

Dengan demikian, turunan MedAct memperlihatkan bahwa gradien tidak pernah menjadi nol secara luas dan tidak mengalami ledakan nilai, menjadikannya kandidat

fungsi aktivasi yang stabil untuk *deep neural network* pada domain citra medis yang kompleks.

C. Analisis Asimtotik

Analisis asimtotik dilakukan untuk memahami perilaku MedAct pada nilai input ekstrem. Ketika $x \rightarrow +\infty$, nilai sigmoid $\sigma(x) \rightarrow 1$, sehingga MedAct mendekati $\text{MedAct} \approx x \cdot \tanh(\alpha)$. Hal ini menunjukkan bahwa MedAct bersifat linear terkontrol pada domain positif.

Sifat linear terkontrol ini mirip dengan ReLU pada domain positif, tetapi dengan faktor skala $\tanh(\alpha)$ yang membatasi amplitudo *output*. Pembatasan ini mencegah pertumbuhan *output* yang terlalu besar yang dapat menyebabkan *exploding gradient* [15].

Sebaliknya, ketika $x \rightarrow -\infty$, maka $\sigma(x) \rightarrow 0$, sehingga $\text{MedAct} \rightarrow 0$. Ini menunjukkan bahwa pada domain negatif, MedAct menghasilkan *output* yang mendekati nol secara halus, bukan secara *abrupt* seperti ReLU.

Perilaku asimtotik ini penting bagi citra medis karena memungkinkan *neuron* untuk tetap memberikan kontribusi kecil pada domain negatif, sehingga informasi dengan intensitas rendah tidak sepenuhnya dihapus dari jaringan. Pendekatan ini sejalan dengan temuan bahwa fungsi aktivasi *smooth* lebih efektif pada data medis dengan distribusi intensitas kompleks [24].

Dengan demikian, MedAct memiliki karakteristik asimtotik yang menggabungkan keunggulan ReLU (linearitas pada domain positif) dan fungsi *smooth* (kelunakan pada domain negatif), sehingga lebih sesuai untuk representasi fitur citra medis.

D. Batas Gradien

Stabilitas gradien merupakan aspek krusial dalam pelatihan *deep neural network*. Untuk MedAct, batas gradien dapat dianalisis menggunakan sifat dasar *sigmoid* dan *tanh*, yaitu $\sigma(x)(1 - \sigma(x)) \leq 0.25$ dan $1 - \tanh^2(z) \leq 1$.

Dengan menggunakan batas tersebut, dapat disimpulkan bahwa kontribusi maksimum dari komponen nonlinier terhadap gradien MedAct juga terbatas. Hal ini memastikan bahwa gradien tidak mengalami lonjakan besar meskipun nilai *input* cukup besar.

Pembatasan gradien ini penting untuk mencegah *exploding gradient*, yang sering menjadi masalah pada jaringan dalam yang dilatih menggunakan fungsi aktivasi *non-bounded*. MedAct, dengan struktur *bounded*-nya, secara teoritis lebih stabil dibanding fungsi aktivasi yang tidak memiliki batas atas eksplisit [1], [7], [9], [10], [11], [13], [32], [33].

Studi terkini menunjukkan bahwa fungsi aktivasi dengan gradien terkontrol menghasilkan konvergensi yang lebih stabil dan lebih cepat, terutama pada dataset medis yang berisik dan tidak seimbang [27], [43].

Oleh karena itu, batas gradien MedAct memberikan justifikasi matematis bahwa fungsi ini aman digunakan pada jaringan yang dalam dan kompleks tanpa meningkatkan risiko ketidakstabilan numerik selama proses pelatihan.

E. Non-Monotonisitas Terbatas

MedAct memiliki sifat non-monotonisitas terbatas akibat interaksi antara komponen linear dan fungsi gating

$$g(x) = \tanh(\alpha \cdot \sigma(x)).$$

Non-monotonisitas ini memungkinkan fungsi menghasilkan respons yang tidak selalu meningkat seiring dengan peningkatan *input*.

Non-monotonisitas telah terbukti bermanfaat dalam fungsi aktivasi modern seperti Swish dan Mish karena memungkinkan jaringan mempelajari representasi fitur yang lebih kompleks dibanding fungsi monoton seperti ReLU [21], [26], [27], [43].

Pada MedAct, non-monotonisitas tidak bersifat ekstrem karena tetap dibatasi oleh fungsi *tanh* dan *sigmoid*. Hal ini menjaga agar respons fungsi tetap stabil dan tidak menghasilkan fluktuasi besar yang sulit dikontrol.

Dalam konteks citra medis, non-monotonisitas terbatas memungkinkan neuron merespons secara selektif terhadap pola tertentu, seperti tepi jaringan atau anomali kecil, tanpa menghilangkan sensitivitas terhadap variasi intensitas rendah [24], [28], [29], [30], [31].

Dengan demikian, sifat non-monotonisitas terbatas MedAct dapat dipandang sebagai mekanisme untuk memperkaya representasi fitur sekaligus mempertahankan stabilitas numerik, yang merupakan kombinasi ideal untuk aplikasi *deep learning* pada citra medis.

V. DISKUSI

MedAct dirancang sebagai fungsi aktivasi yang menggabungkan keunggulan fungsi linear dan fungsi aktivasi *smooth* dalam satu formulasi matematis yang terpadu. Dengan struktur $MedAct(x) = x \cdot \tanh(\alpha \cdot \sigma(x))$, MedAct mempertahankan komponen linear yang penting untuk propagasi informasi sekaligus menyediakan mekanisme nonlinier yang halus melalui kombinasi *sigmoid* dan *tanh*. Pendekatan ini menempatkan MedAct sebagai alternatif konseptual terhadap fungsi aktivasi konvensional yang bersifat *piecewise* atau tidak terdiferensialkan secara penuh.

Jika dibandingkan dengan ReLU, keunggulan utama MedAct terletak pada kemampuannya untuk tetap menghasilkan keluaran non-nol pada domain negatif secara halus. ReLU memotong seluruh nilai negatif menjadi nol, yang berpotensi menghilangkan informasi penting terutama pada citra medis dengan perbedaan intensitas kecil antar jaringan. MedAct, melalui mekanisme *gating* adaptif, memungkinkan kontribusi nilai negatif tetap dipertahankan dalam bentuk terkontrol sehingga mengurangi risiko terjadinya fenomena *dead neuron*.

Dibandingkan dengan fungsi aktivasi modern seperti Swish dan Mish, MedAct menawarkan kontrol tambahan melalui parameter α (alpha). Parameter ini berfungsi sebagai pengatur tingkat nonlinieritas sehingga respons fungsi dapat disesuaikan dengan karakteristik data yang dihadapi. Dengan demikian, MedAct tidak hanya bersifat *smooth* dan non-monotonik terbatas, tetapi juga adaptif

terhadap distribusi *input*, suatu sifat yang tidak secara eksplisit tersedia pada Swish dan Mish tanpa parameter tambahan.

Dalam konteks citra medis, potensi manfaat MedAct menjadi lebih signifikan. Citra medis sering mengandung informasi penting pada rentang intensitas rendah, misalnya pada batas jaringan atau lesi kecil. Fungsi aktivasi yang terlalu agresif dalam memotong nilai negatif atau mendistorsi variasi kecil dapat menurunkan kualitas representasi fitur. MedAct berpotensi mempertahankan informasi intensitas rendah sekaligus meningkatkan sensitivitas jaringan terhadap variasi morfologi yang relevan secara klinis.

Selain itu, sifat kontinu, terdiferensialkan penuh, dan memiliki batas gradien yang stabil menjadikan MedAct secara teoritis lebih aman digunakan pada jaringan saraf dalam yang kompleks. Stabilitas ini penting untuk menjaga konvergensi pelatihan dan mengurangi fluktuasi gradien yang dapat menghambat proses optimasi. Oleh karena itu, MedAct dapat dipandang sebagai fungsi aktivasi yang secara konseptual sesuai untuk aplikasi *deep learning* pada citra medis, dengan keseimbangan antara fleksibilitas representasi dan stabilitas numerik.

VI. KETERBATASAN DAN ARAH PENELITIAN LANJUTAN

Penelitian ini bersifat teoretis dan berfokus pada perancangan serta analisis matematis fungsi aktivasi MedAct tanpa melibatkan pengujian empiris pada *dataset* nyata. Oleh karena itu, kesimpulan yang diperoleh masih berada pada ranah konseptual dan belum dapat secara langsung menggambarkan performa MedAct dalam skenario praktis. Ketiadaan validasi eksperimental menjadi keterbatasan utama yang perlu diperhatikan dalam interpretasi hasil kajian ini.

Penelitian lanjutan perlu menguji MedAct pada berbagai arsitektur *deep learning*, khususnya *Convolutional Neural Network* (CNN) dan *Residual Network* (ResNet), yang umum digunakan dalam analisis citra medis. Implementasi MedAct sebagai pengganti fungsi aktivasi standar seperti ReLU akan memungkinkan evaluasi langsung terhadap pengaruhnya pada akurasi klasifikasi, kualitas segmentasi, serta stabilitas pelatihan jaringan.

Selain itu, MedAct perlu dibandingkan secara sistematis dengan fungsi aktivasi populer seperti ReLU, Swish, dan Mish. Perbandingan ini harus dilakukan menggunakan metrik evaluasi yang konsisten, seperti akurasi, presisi, recall, F1-score, serta kurva konvergensi pelatihan. Pendekatan komparatif ini penting untuk menilai apakah keunggulan teoretis MedAct benar-benar terwujud dalam performa empiris.

Aspek penting lain yang perlu dikaji adalah pengaruh parameter α (alpha) terhadap dinamika pembelajaran. Eksperimen dapat dilakukan dengan membandingkan varian *Fixed- α* dan *Learnable- α* untuk mengetahui sejauh mana fleksibilitas parameter ini meningkatkan kemampuan representasi jaringan atau justru menimbulkan instabilitas

gradien. Analisis sensitivitas terhadap nilai α (alpha) juga diperlukan untuk menentukan rentang nilai yang optimal dan stabil.

Lebih lanjut, penelitian lanjutan perlu mengkaji stabilitas konvergensi dan perilaku gradien MedAct selama proses pelatihan jaringan dalam. Hal ini mencakup analisis kecepatan konvergensi, fluktuasi *loss*, serta ketahanan terhadap *noise* dan ketidakseimbangan data. Dengan melakukan kajian empiris yang komprehensif, MedAct dapat dievaluasi secara objektif sebagai kandidat fungsi aktivasi baru yang layak digunakan dalam aplikasi *deep learning* pada citra medis.

VII. KESIMPULAN

Artikel ini mengusulkan MedAct sebagai fungsi aktivasi baru yang dirancang secara teoretis melalui formulasi hibrida antara komponen linear, sigmoid, dan fungsi hiperbolik. MedAct dikembangkan dengan mempertimbangkan kebutuhan khusus domain citra medis yang menuntut stabilitas numerik, sensitivitas terhadap variasi intensitas kecil, serta kemampuan representasi nonlinier yang adaptif. Pendekatan ini menempatkan MedAct sebagai kontribusi konseptual dalam pengembangan fungsi aktivasi berbasis domain (*domain-specific activation function*) untuk *deep learning* medis.

Analisis matematis yang dilakukan menunjukkan bahwa MedAct bersifat kontinu dan terdiferensialkan pada seluruh domain *real*, memiliki batas gradien yang stabil, serta menunjukkan perilaku asimtotik yang terkontrol. Sifat-sifat ini mengindikasikan bahwa MedAct secara teoritis aman digunakan pada *deep neural network* tanpa meningkatkan risiko *vanishing gradient* maupun *exploding gradient*. Selain itu, karakter non-monotonisitas terbatas yang dimiliki MedAct memungkinkan representasi fitur yang lebih kaya dibanding fungsi aktivasi monoton seperti ReLU.

Dengan demikian, MedAct diharapkan dapat menjadi landasan konseptual bagi penelitian eksperimental selanjutnya yang menguji efektivitasnya pada arsitektur CNN dan ResNet untuk berbagai tugas citra medis. Kontribusi utama artikel ini terletak pada penyediaan justifikasi matematis yang kuat terhadap desain fungsi aktivasi baru, sehingga dapat memperkaya kajian ilmiah mengenai peran fungsi aktivasi dalam *deep learning* medis dan membuka peluang pengembangan fungsi aktivasi berbasis karakteristik domain secara lebih sistematis di masa depan.

DAFTAR PUSTAKA

- [1] G. Litjens *et al.*, "A survey on deep learning in medical image analysis.," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [2] S. ATASEVER, N. AZGINOGLU, D. S. TERZI, and R. TERZI, "A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning," *Clin. Imaging*, vol. 94, pp. 18–41, 2023, doi: https://doi.org/10.1016/j.clinimag.2022.11.003.
- [3] I. D. Apostolopoulos, D. I. Apostolopoulos, T. I. Spyridonidis, N. D. Papathanasiou, and G. S. Panayiotakis, "Multi-input deep learning approach for Cardiovascular Disease diagnosis using Myocardial Perfusion Imaging and clinical data," *Phys. Medica Eur. J. Med. Phys.*, vol. 84, pp. 168–177, Apr. 2021, doi: 10.1016/j.ejmp.2021.04.011.
- [4] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation.," *IEEE Trans. Med. Imaging*, vol. 39, no. 6, pp. 1856–1867, June 2020, doi: 10.1109/TMI.2019.2959609.
- [5] D. Abdelraouf, M. Essam, and M. Elattar, "Light-Weight Localization and Scale-Independent Multi-gate UNET Segmentation of Left and Right Ventricles in MRI Images," *Cardiovasc. Eng. Technol.*, vol. 13, no. 3, pp. 393–406, June 2022, doi: 10.1007/s13239-021-00591-2.
- [6] C. L. Angelina and A. Rospawan, "A benchmark study of deeplabv3+, u-net++, and attention u-net for blood cell segmentation," *Green Intell. Syst. Appl.*, vol. 5, no. 1, pp. 61–73, 2025.
- [7] G. Wu and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, Mar. 2017, doi: 10.1146/annurev-bioeng-071516-044442.
- [8] X. Liu *et al.*, "Advances in Deep Learning-Based Medical Image Analysis.," *Health Data Sci.*, vol. 2021, p. 8786793, 2021, doi: 10.34133/2021/8786793.
- [9] L. Fereid and J. Mathew, "Deep Learning in Radiology: Current Capabilities and Future Directions in Diagnostic Imaging," 2025.
- [10] Y. Wang, X. Ge, H. Ma, S. Qi, G. Zhang, and Y. Yao, "Deep Learning in Medical Ultrasound Image Analysis: A Review," *IEEE Access*, vol. PP, pp. 1–1, Apr. 2021, doi: 10.1109/ACCESS.2021.3071301.
- [11] J. Wang, S. Wang, and Y. Zhang, "Deep learning on medical image analysis," *CAAI Trans. Intell. Technol.*, vol. 10, pp. 1–35, June 2024, doi: 10.1049/cit2.12356.
- [12] A. Meca, "Applications of Deep Learning to Magnetic Resonance Imaging (MRI)," in *2023 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, Los Alamitos, CA, USA: IEEE Computer Society, Aug. 2023, pp. 113–120. doi: 10.1109/iCCECE59400.2023.10238598.
- [13] P. M. Cheng *et al.*, "Deep Learning: An Update for Radiologists.," *Radiogr. Rev. Publ. Radiol. Soc. N. Am. Inc.*, vol. 41, no. 5, pp. 1427–1445, Oct. 2021, doi: 10.1148/rg.2021200210.
- [14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [15] C. C. Aggarwal, "Neural Networks," in *Artificial Intelligence: A Textbook*, Cham: Springer International Publishing, 2021, pp. 211–251. doi: 10.1007/978-3-030-72357-6_7.
- [16] Y. Ma and J. Tang, *Deep Learning on Graphs*. Cambridge University Press, 2021.
- [17] A. Apicella, F. Donnarumma, F. Isgrò, and R. Prevete, "A survey on modern trainable activation functions," *Neural Netw.*, vol. 138, pp. 14–32, 2021, doi: https://doi.org/10.1016/j.neunet.2021.01.026.
- [18] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation.," *Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support 4th Int. Workshop DLMIA 2018 8th Int. Workshop ML-CDS 2018 Held Conjunction MICCAI 2018 Granada Spain S*, vol. 11045, pp. 3–11, Sept. 2018, doi: 10.1007/978-3-030-00889-5_1.
- [19] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nat. Med.*, vol. 25, Jan. 2019, doi: 10.1038/s41591-018-0316-z.
- [20] S. K. Zhou *et al.*, "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises.," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 109, no. 5, pp. 820–838, May 2021, doi: 10.1109/JPROC.2021.3054390.
- [21] D. Misra, "Mish: A Self Regularized Non-Monotonic Activation Function." 2020. [Online]. Available: https://arxiv.org/abs/1908.08681
- [22] S. Elfwing, E. Uchibe, and K. Doya, "Sigmoid-Weighted Linear Units for Neural Network Function Approximation in Reinforcement Learning." 2017. [Online]. Available: https://arxiv.org/abs/1702.03118
- [23] A. Paul, R. Bandyopadhyay, J. H. Yoon, Z. W. Geem, and R. Sarkar, "SinLU: Sinu-Sigmoidal Linear Unit," *Mathematics*, vol. 10, no. 3, 2022, doi: 10.3390/math10030337.

- [24] S. Khan, H. Rahmani, {Syed Afaq Ali} Shah, and M. Bennamoun, *A Guide to Convolutional Neural Networks for Computer Vision*. in Synthesis Lectures on Computer Vision, no. 1. Morgan & Claypool Publishers, 2018. doi: 10.2200/S00822ED1V01Y201712COV015.
- [25] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: a survey.," *Evol. Intell.*, vol. 15, no. 1, pp. 1–22, 2022, doi: 10.1007/s12065-020-00540-3.
- [26] A. Rahman, L. He, and H. Wang, "Activation function optimization scheme for image classification," *Knowl.-Based Syst.*, vol. 305, p. 112502, 2024, doi: <https://doi.org/10.1016/j.knosys.2024.112502>.
- [27] K. Biswas *et al.*, "Adaptive Smooth Activation Function for Improved Organ Segmentation and Disease Diagnosis," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, M. G. Linguraru, Q. Dou, A. Feragen, S. Giannarou, B. Glocker, K. Lekadir, and J. A. Schnabel, Eds., Cham: Springer Nature Switzerland, 2024, pp. 65–74.
- [28] L. Szilagyi and L. Kovács, "Special Issue: Artificial Intelligence Technology in Medical Image Analysis," *Appl. Sci.*, vol. 14, p. 2180, Mar. 2024, doi: 10.3390/app14052180.
- [29] S. Khan *et al.*, "Advanced Feature Selection Techniques in Medical Imaging—A Systematic Literature Review," *Comput. Mater. Contin.*, vol. 85, no. 2, pp. 2347–2401, 2025, doi: <https://doi.org/10.32604/cmc.2025.066932>.
- [30] Y. Bian, J. Li, C. Ye, X. Jia, and Q. Yang, "Artificial intelligence in medical imaging: From task-specific models to large-scale foundation models.," *Chin. Med. J. (Engl.)*, vol. 138, no. 6, pp. 651–663, Mar. 2025, doi: 10.1097/CM9.0000000000003489.
- [31] A. Choudhary, L. Tong, Y. Zhu, and M. D. Wang, "Advancing Medical Imaging Informatics by Deep Learning-Based Domain Adaptation.," *Yearb. Med. Inform.*, vol. 29, no. 1, pp. 129–138, Aug. 2020, doi: 10.1055/s-0040-1702009.
- [32] S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, "Activation Functions in Deep Learning: A Comprehensive Survey and Benchmark." 2022. [Online]. Available: <https://arxiv.org/abs/2109.14545>
- [33] A. Primbetov and N. Akbarov, "A COMPARATIVE STUDY OF ACTIVATION FUNCTIONS IN DEEP LEARNING MODELS," *Al-Fargoniy Avlodlari*, vol. 1, no. 3, pp. 80–84, Oct. 2025, doi: 10.5281/zenodo.17258843.
- [34] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)." 2016. [Online]. Available: <https://arxiv.org/abs/1511.07289>
- [35] A. A. Yahya, M. Al-Qatf, and J. Jiang, "Balanced Learnable Activation Function (BLAF): Enhancing Classification Accuracy in Deep Neural Networks," *J. Intell. Fuzzy Syst.*, vol. 49, no. 5, pp. 1265–1290, 2025, doi: 10.1177/18758967251343172.
- [36] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-Normalizing Neural Networks." 2017. [Online]. Available: <https://arxiv.org/abs/1706.02515>
- [37] Z. Chen *et al.*, "Large-scale self-normalizing neural networks," *J. Autom. Intell.*, vol. 3, no. 2, pp. 101–110, 2024, doi: <https://doi.org/10.1016/j.jai.2024.05.001>.
- [38] Z. Wang, H. Li, and L. Ma, "Modern synergetic neural network for imbalanced small data classification.," *Sci. Rep.*, vol. 13, no. 1, p. 15669, Sept. 2023, doi: 10.1038/s41598-023-42689-8.
- [39] A. Waqas *et al.*, "Self-Normalizing Multi-Omics Neural Network for Pan-Cancer Prognostication.," *Int. J. Mol. Sci.*, vol. 26, no. 15, July 2025, doi: 10.3390/ijms26157358.
- [40] K. Al-Hammuri, F. Gebali, A. Kanan, and I. T. Chelvan, "Vision transformer architecture and applications in digital health: a tutorial and survey.," *Vis. Comput. Ind. Biomed. Art*, vol. 6, no. 1, p. 14, July 2023, doi: 10.1186/s42492-023-00140-9.
- [41] X. Jiang, Z. Hu, S. Wang, and Y. Zhang, "Deep Learning for Medical Image-Based Cancer Diagnosis.," *Cancers*, vol. 15, no. 14, July 2023, doi: 10.3390/cancers15143608.
- [42] S. Khan, H. Rahmani, S. Shah, and M. Bennamoun, "A Guide to Convolutional Neural Networks for Computer Vision," *Synth. Lect. Comput. Vis.*, vol. 8, pp. 1–207, Feb. 2018, doi: 10.2200/S00822ED1V01Y201712COV015.
- [43] P. Ren, T. Pan, G. Yang, Y. Guo, W. Wei, and Z. Pan, "ReLU-like non-monotonic smooth activation functions based on regularized heaviside functions and extensions," *Math. Found. Comput.*, vol. 9, Jan. 2025, doi: 10.3934/mfc.2025009.
- [44] K. Biswas, M. Karri, and U. Bağcı, "A Non-monotonic Smooth Activation Function." 2023. [Online]. Available: <https://arxiv.org/abs/2310.10126>