

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Priority Mail Express® Label No. _____

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Additional inventors are being named on the _____ separately numbered sheets attached hereto.		
TITLE OF THE INVENTION (500 characters max):		
Direct all correspondence to: CORRESPONDENCE ADDRESS <input type="checkbox"/> The address corresponding to Customer Number: <div style="border: 1px solid black; width: 250px; height: 25px; display: inline-block; vertical-align: middle;"></div>		
OR <input type="checkbox"/> Firm or Individual Name		
Address		
City	State	Zip
Country	Telephone	Email
ENCLOSED APPLICATION PARTS (check all that apply)		
<input type="checkbox"/> Application Data Sheet. See 37 CFR 1.76.		
<input type="checkbox"/> CD(s), Number of CDs _____		
<input type="checkbox"/> Drawing(s) Number of Sheets _____		
<input type="checkbox"/> Other (specify) _____		
<input type="checkbox"/> Specification (e.g., description of the invention) Number of Pages _____		
Fees Due: Filing Fee of \$325 (\$130 for small entity) (\$65 for micro entity). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is \$450 (\$180 for small entity) (\$90 for micro entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).		
METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT		
<input type="checkbox"/> Applicant asserts small entity status. See 37 CFR 1.27.		
<input type="checkbox"/> Applicant certifies micro entity status. See 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.		
<input type="checkbox"/> A check or money order made payable to the <i>Director of the United States Patent and Trademark Office</i> is enclosed to cover the filing fee and application size fee (if applicable).		<div style="border: 1px solid black; width: 100px; height: 25px; margin-bottom: 5px;"></div> TOTAL FEE AMOUNT (\$)
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.		
<input type="checkbox"/> The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: _____		

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

A Federal agency may not conduct or sponsor, and a person is not required to respond to, nor shall a person be subject to a penalty for failure to comply with an information collection subject to the requirements of the Paperwork Reduction Act of 1995, unless the information collection has a currently valid OMB Control Number. The OMB Control Number for this information collection is 0651-0032. Public burden for this form is estimated to average 18 hours per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the information collection. Send comments regarding this burden estimate or any other aspect of this information collection, including suggestions for reducing this burden to the Chief Administrative Officer, United States Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450 or email InformationCollection@uspto.gov. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **If filing this completed form by mail, send to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 2 of 2

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government. (NOTE: Providing this information on a provisional cover sheet, such as this Provisional Application for Patent Cover Sheet (Form PTO/SB/16), does not satisfy the requirement of 35 U.S.C. 202(c)(6), which requires that the *specification* contain a statement specifying that the invention was made with Government support and that the Government has certain rights in the invention.)

☐ No.

☐ Yes, the invention was made by an agency of the U.S. Government. The U.S. Government agency name is:

☐ Yes, the invention was made under a contract with an agency of the U.S. Government.

The contract number is: _____

The U.S. Government agency name is: _____

In accordance with 35 U.S.C. 202(c)(6) and 37 CFR 401.14(f)(4), the specifications of any United States patent applications and any patent issuing thereon covering the invention, including the enclosed provisional application, must state the following:

“This invention was made with government support under [IDENTIFY THE CONTRACT] awarded by [IDENTIFY THE FEDERAL AGENCY]. The government has certain rights in the invention.”

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

SIGNATURE _____ DATE _____

TYPED OR PRINTED NAME _____ REGISTRATION NO. _____
(if appropriate)

TELEPHONE _____ DOCKET NUMBER _____

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. The United States Patent and Trademark Office (USPTO) collects the information in this record under authority of 35 U.S.C. 2. The USPTO's system of records is used to manage all applicant and owner information including name, citizenship, residence, post office address, and other information with respect to inventors and their legal representatives pertaining to the applicant's/owner's activities in connection with the invention for which a patent is sought or has been granted. The applicable Privacy Act System of Records Notice for the information collected in this form is COMMERCE/PAT-TM-7 Patent Application Files, available in the Federal Register at 78 FR 19243 (March 29, 2013). <https://www.govinfo.gov/content/pkg/FR-2013-03-29/pdf/2013-07341.pdf>

Routine uses of the information in this record may include disclosure to:

- 1) law enforcement, in the event that the system of records indicates a violation or potential violation of law;
- 2) a Federal, state, local, or international agency, in response to its request;
- 3) a contractor of the USPTO having need for the information in order to perform a contract;
- 4) the Department of Justice for determination of whether the Freedom of Information Act (FOIA) requires disclosure of the record;
- 5) a Member of Congress submitting a request involving an individual to whom the record pertains, when the individual has requested the Member's assistance with respect to the subject matter of the record;
- 6) a court, magistrate, or administrative tribunal, in the course of presenting evidence, including disclosures to opposing counsel in the course of settlement negotiations;
- 7) the Administrator, General Services Administration (GSA), or their designee, during an inspection of records conducted by GSA under authority of 44 U.S.C. 2904 and 2906, in accordance with the GSA regulations and any other relevant (i.e., GSA or Commerce) directive, where such disclosure shall not be used to make determinations about individuals;
- 8) another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c));
- 9) the Office of Personnel Management (OPM) for personnel research purposes; and
- 10) the Office of Management and Budget (OMB) for legislative coordination and clearance.

If you do not furnish the information requested on this form, the USPTO may not be able to process and/or examine your submission, which may result in termination of proceedings, abandonment of the application, and/or expiration of the patent.

Title: *Systems and Methods for Cancer Classification Using Multi-Modal Transformer-Based Architectures*

Inventor: Dr. R. Craig Stillwell

Filing Date: July 20, 2025

1. FIELD OF THE INVENTION

This invention relates to the field of biomedical informatics and machine learning. Specifically, it describes a system and method for the classification of cancer types using multi-modal genomic data integrated through transformer-based artificial intelligence architectures.

2. BACKGROUND OF THE INVENTION

Accurate classification of cancer types is essential for diagnosis, treatment planning, and personalized medicine. Although multiple high-throughput biological datasets are now available — including transcriptomics (e.g., RNA-seq), DNA methylation, and histone modification signals — few AI systems leverage these diverse modalities in a unified, explainable, and production-ready way.

Traditional diagnostic models either use unimodal data or are research prototypes that lack deployment infrastructure, scalability, or clinical interpretability. This invention addresses that gap.

3. SUMMARY OF THE INVENTION

The invention, referred to as **Cancer Alpha**, is a production-ready, multi-modal AI system for cancer classification. It uses multiple sources of biological data, processed through domain-specific encoders, fused using a transformer-based architecture, and passed through an ensemble classifier for high-accuracy predictions. The system includes an interpretability layer to produce feature-level importance rankings and a fully containerized infrastructure for deployment in clinical or research settings.

Cancer Alpha is deployed using FastAPI, containerized via Docker and Kubernetes, monitored using Prometheus and Grafana, and integrated with secure APIs for clinical usage.

4. DETAILED DESCRIPTION OF THE SYSTEM

4.1 Data Sources

Cancer Alpha ingests and harmonizes data from:

- Gene expression datasets (e.g., TCGA RNA-seq)
- DNA methylation data (e.g., Illumina arrays)
- Histone modification profiles (e.g., ENCODE ChIP-seq)

4.2 Preprocessing Pipeline

- Each modality is independently normalized, scaled, and filtered.
- Missing values are imputed using modality-specific techniques.
- Dimensionality reduction is applied (e.g., PCA, feature selection).

4.3 Architecture

- **Modality-specific encoders** extract latent feature representations for each data type.
- A **multi-modal transformer fusion layer** performs cross-modal learning using self- and cross-attention.
- The fused embeddings are passed to an **ensemble classification layer** (e.g., Random Forest, XGBoost) that outputs the predicted cancer type.
- An **attention-based interpretability layer** ranks the biological features contributing to the decision.

4.4 Output

- Predicted cancer classification (e.g., lung, breast, colon)
- Ranked list of contributing features (genes, methylation sites, histone markers)

4.5 Deployment Infrastructure

- FastAPI-based backend for real-time inference
- REST API endpoints for integration with lab systems or diagnostic interfaces
- Containerization via Docker; orchestration via Kubernetes
- Prometheus and Grafana for real-time monitoring and system observability
- Redis cache layer for inference speed and efficiency
- CI/CD pipeline implemented via GitHub Actions for reproducible deployment

5. USE CASES

- Early diagnostic support for oncologists using minimally invasive samples
- Multi-omics research platforms exploring biomarkers and pathway analysis
- Integration into lab-based or hospital-based clinical workflows
- Cloud-based cancer inference services for precision oncology

6. LEGAL NOTES

This application is a **provisional patent application** and does not include formal claims. However, the invention disclosed herein is intended to establish intellectual property priority for:

- The integration of multi-modal omics via a transformer-based AI architecture
- The inclusion of an explainability layer for feature attribution
- The containerized deployment of said model using real-time APIs and monitoring tools