MARCH 2022 ISSUE 02

AI in HEALTHCARE

MARKTECHPOST

2022

Top Innovative AI Startups in Healthcare

MARKTECHPOST

Exclusive Interview with PATRICK BANGERT
Vice President of AI at Samsung SDS

Exclusive Interview with KIMBERLY POWELL
VP and General Manager of NVIDIA Healthcare

An Italian AI Startup Turning Emotional Unconsciousness To Instant Music'
By: INNEREO TEAM (Italy)

In Afghanistan Women Are Fighting A War Against An Invisible Enemy – Breast Cancer!
By: Tarry Singh

EDITOR'S LETTER

By: Asif Razzaq

In today's fast-paced world where technology changes rapidly and new innovations come out every day, artificial intelligence (AI) is revolutionizing how we live, making its way into every industry and transforming the way people work. AI has given rise to new innovations such as music recommendations in the entertainment sectors or computer vision impacting transportation models - it's even helping with early diagnoses for diseases like cancer. Therefore, we are bringing a magazine series, Marktechpost AI Magazine, featuring the latest applications of AI in various domains.

The Marktechpost AI Magazine is back with a second edition that will explore the latest developments for AI in healthcare technology. From voice-driven assistants to automated diagnosis, this issue has it all. We'll be featuring interesting articles about how artificial intelligence can make your life easier at home or work as well interviews highlighting what businesses need for success now more than ever before.

Marktechpost is an AI News Platform providing easy-to-consume, byte size updates in machine learning,deep learning and data science research. Our vision is to showcase the hottest research trends in AI from around the world using our innovative method of search and discovery.

We welcome feedback from the AI community! Share your thoughts at:
asif@marktechpost.com

FIRST ISSUE

Published by Marktechpost, LLC

EDITORIAL ADVISORY BOARD

MARKTECHPOST TEAM

GUEST EDITORS

Tarry Singh:

Tarry Singh is Chairman, CEO, and AI Researcher of AI company Real AI Inc and deepkapha.ai. Tarry has 25 years of experience working with data and has advised CxOs of global organizations to set up data-driven organizations from scratch. He speaks regularly at global AI leadership summits worldwide and conducts workshops on a regular basis with his TAs who are currently PhDs in various disciplines such as NLP, Computer Vision, Robotics, and other Artificial Intelligence disciplines. He is a board member at HU University Utrecht for AI/Machine Learning as well as Visiting Faculty for AI at the University of Texas Dallas. Tarry has also been an AI Mentor at Coursera, the world's #2 Specialization according to Inc.

Tarry writes in Forbes Magazine about the business impact of AI, he was chosen as the top 10 "LinkedIn Voice for DataScience & Analytics 2018" from 600M professionals worldwide and he has been interviewed by leading press such as Al Jazeera/Associated Press to name a few. Since then, he has consistently being chosen as the leading AI expert alongside Alan Turing AI laureates.

Jean-Marc Mommessin:

Jean-Marc is a successful AI business executive. He leads and accelerates growth for AI powered solutions and started a computer vision company in 2006. He is a recognized speaker at AI conferences and has an MBA from Stanford.

Table of Contents

Exclusive Interview with

PATRICK BANGERT

VP of AI at Samsung SDS

BIO:

Patrick heads the AI Division at Samsung SDSA. On the side of AI Engineering, he is responsible for Brightics AI Accelerator, a distributed ML training and automated ML product, and AutoLabel, an automatic image data annotation and modeling tool primarily targeted at the medical imaging community. On the side of AI Sciences, he leads the consulting group that makes AI models for our customers for many diverse use cases. Among his other responsibilities is to act as a visionary for the future of AI at Samsung.

Before joining Samsung, Patrick spent 15 years as CEO at Algorithmica Technologies, a machine learning software company serving the chemicals and oil and gas industries. Prior to that, he was an assistant professor of applied mathematics at Jacobs University in Germany, as well as a researcher at Los Alamos National Laboratory and NASA's Jet Propulsion Laboratory. Patrick obtained his machine learning PhD in mathematics and his Masters's in theoretical physics from University College London.

Can you share about your journey in AI so far?

As a mathematician and physicist, I've always been interested in how to generate accurate mathematical descriptions of the world. AI allows us to do that quickly, accurately, and with realistic effort. After getting my PhD, Jacobs University in Germany offered me a professorship in applied mathematics. I worked as a professor for three years prior to starting my own company. Algorithmica technologies deals with AI applications to the oil and gas and chemicals industries. The primary questions are predictive maintenance and process optimization where AI is used to model physical reality such as the mechanics of machines or chemical reactions. For the past two years, I lead the AI division of Samsung SDS in San Jose, California.

Tell us about your role as VP of AI at Samsung SDS. What are some of the biggest challenges around leading and directing AI teams?

The AI division produces software to help our customers make AI models, and we also make models ourselves in a consultative capacity. Our software automates the AI workflow to a very large degree including data annotation, feature engineering, model selection, and hyper-parameter tuning. The engineering (software), science (modeling service), and sales and marketing teams are part of the division and we operate both within Samsung Group as well as for third-party customers. The biggest challenge with AI teams is hiring and retention as AI is a very popular area with many more jobs than qualified candidates. Our greatest resource is great minds who innovate and creatively solve problems. Another challenge is aligning all the details, people, and processes behind a concrete and simple business goal.

Tell us about Samsung Data Services (SDS). How is Samsung delivering AI solutions to its customers?

There are many internal applications of AI in the manufacturing parts of Samsung Group where we try to find and avoid scrap production. The logistics of transporting goods from manufacturing to warehouses are planned. Distribution to stores in light of what AI says is going to be bought at each of a great many international stores during which week is dealt with as well. Then there are external applications that are visible to users such as the Bixby service on the phone. Samsung SDS operates numerous data centers around the world that provide the cloud infrastructure and service to provide owners of Samsung products with the required services. Many AI models are examples of this. Our latest effort is in healthcare where we analyze medical images to diagnose diseases by AI.

Can you shed some light on Brightics AI Accelerator?

The Brightics AI Accelerator (AIA) is our flagship product that provides automated AI/ML modeling capabilities to anyone. It is a cloud service that can take you from raw dataset to a final and good model in the shortest possible time. First, it shortens the data annotation effort by 90% with its product offering AutoLabel, which is crucial as this is the most time and effort-consuming process in all of AI. Second, it transforms the dataset by feature engineering into a form that better represents the information content of the data. Third, it selects the right model that is the best fit for the data. Fourth, it tunes the hyper-parameters of the training algorithm so that the model can be optimally fitted. Fifth, it does the actual modeling in a distributed way so as to use many computers simultaneously to train a single model thereby saving a lot of time. In this way, we can collapse a project of a year into a couple of weeks and reduce model development budgets by millions of dollars.

What are the primary issues in healthcare that AI can help to solve?

Financially, annotating the raw dataset is the fulcrum of AI. Our product AutoLabel, as part of AIA, solves this problem. We are focusing on healthcare imaging in our effort to create AI that helps us all. Image AI models are 99% accurate as compared to 70% for human doctors. These models can provide instant feedback whereas hospital pathology workflows can take 4-6 weeks to return a result. The treatment can begin that much earlier and thus provide better outcomes and save lives. Such models live in the cloud and can be accessed by anyone, at any time, from any place in the world. The triple benefits of automation, rapidity, and cost-effectiveness means that imaging can occur more often, and also in a preventative capacity, so as to maintain your current good health instead of just reacting to an already worrisome health threat. Finally, this can help underserved areas of the world that may have equipment but not enough qualified medical staff to reliably interpret the images.

A most interesting exception to healthcare imaging that usually requires expensive equipment is the skin. The skin can be photographed by yourself using your (Samsung) mobile phone. The image can then be analyzed instantly by a cloud service for skin conditions or even aesthetic assessment from a cosmetic point of view.

Please give us an overview of the various healthcare applications in which Samsung AI has been applied.

Our models can diagnose Covid-19 based on an x-ray image of your lungs. They can detect breast cancer based on a microscope image of a biopsy. They can identify the correct place to put a stent into your veins to prevent heart attacks based on intravascular ultrasound or analyze the structure of your veins based on angiography. They can find colon cancer based on a colonoscopy video. They can reconstruct the 3D structure of a brain tumor based on several 2D MRI images slices of your brain. They can diagnose a variety of skin ailments based on a selfie. These are some examples of what can be done. In all applications, the model has a very high accuracy easily beating any human expert – and doing this instantly with negligible cost.

What has been your experience in using AI to predict serious illnesses? How has the Samsung AI system met with success in this sphere?

Our experience so far is scientific in the sense that we've demonstrated that AI can reliably detect serious illnesses. The benefit to the doctor and hospital system is that this saves time and cost. It enables the system to spend more time with the patient rather than with the data. The benefit to the patient is a more reliable and robust diagnosis, an earlier start of treatment, and thus a better chance for a good outcome. Samsung SDS is working with medical device companies and hospital systems to deploy these models either by embedding them in devices, or serving them via the cloud.

What do you foresee as the biggest trends in AI in 2022?

Autonomous driving is the most famous use case but essentially a solved one, from an AI perspective. While a full level 5 car is still in the future, almost all normal driving scenarios don't require that. I consider the driving case a done deal that is now just waiting to be rolled out across the world. Retail has already heavily adopted AI in many processes. Government poses a large opportunity but this is also fraught with many difficulties so I don't see this coming in the next few years. Healthcare is the next big frontier for AI. Prepare to be met with a hybrid system in the next few years where doctors are able to spend time with you and actually look at you during a consultation. It will be AI that takes the notes, prescribes the medicine, and analyzes the data. The doctor will check all this, make the decisions, but ultimately take better care of you on a human level.

AN INTRODUCTION TO AI FOR HEALTHCARE

INTRODUCTION

Artificial Intelligence (AI) and its use are becoming a reality in many medical fields and specialties. Whether it's being used to personalize treatment for a patient, power surgical robots that assist surgeons, or even facilitate back-end operations in hospitals, AI has been a great asset to the healthcare industry.

But first, it's worth asking oneself - what is AI? ***The European Parliament*** defines AI as "the capability of a computer program to perform tasks or reasoning processes that we usually associate with intelligence in a human

being." AI can lead to better care outcomes and improve the productivity and efficiency of care delivery. It can also improve the day-to-day life of healthcare practitioners. ***McKinsey reports*** that letting doctors spend more time looking after patients raises staff morale and improves retention. It can even get life-saving treatments to market faster.

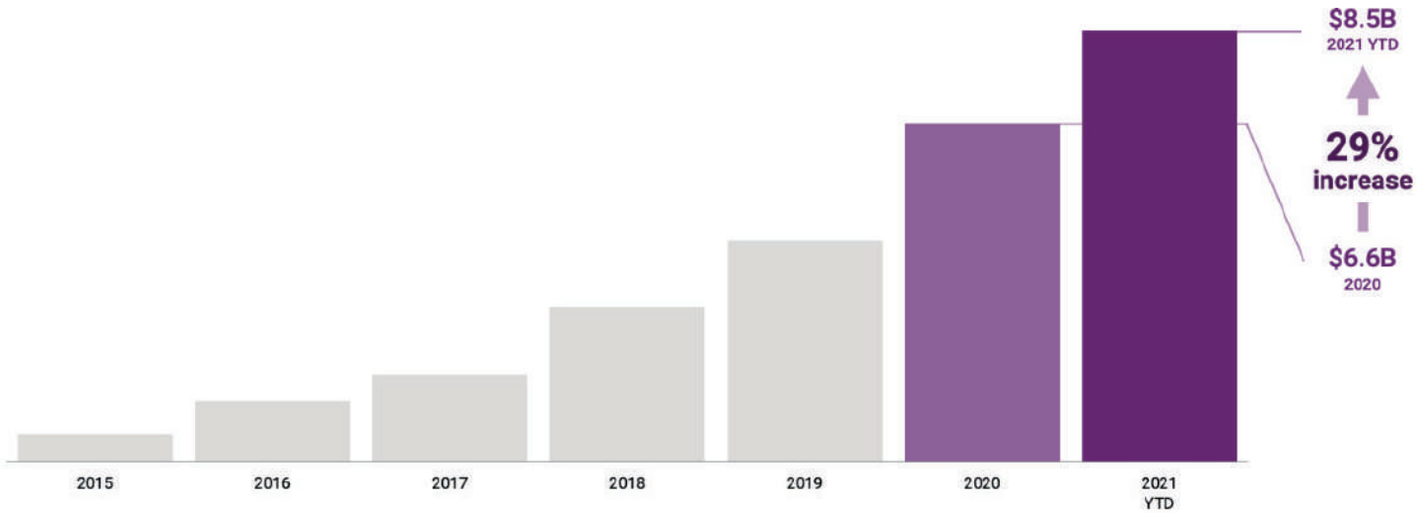
Other tools like machine learning (ML), natural language processing (NLP), and deep learning (DL) enable different stakeholders to identify hidden patterns, assess healthcare needs and solutions faster with more accuracy, informing medical and business decisions across the value chain. AI algorithms parse data sets to identify and label data patterns, while NLP allows these algorithms to isolate relevant data. With DL, the data is analyzed and interpreted.

The market has grown tremendously: Frost & Sullivan ***reported*** that the market was $630 million in 2014. According to ***Allied Market Research***, the market was valued at $8.23 billion in 2020 and is projected to reach $194.4 billion by 2030.

The pandemic led to an increase in investments in healthcare artificial intelligence, with $8.5 billion being raised globally in 2021, according to a ***CBInsights*** report published on November 24th, 2021. The funding figures echo an upward trend: healthcare AI startups raised a record $6.6 billion in 2020, but this figure was surpassed in 2021, with 366 deals done this year to date. The U.S. led global funding of healthcare AI in the third quarter with $2.1 billion, compared to Asia's $800 million and Europe's $100 million.

While optimistic about the role that AI will play in healthcare, questions have been raised about the impact it could have on patients, practitioners, and health systems, and about its potential risks. There are ethical concerns around how AI and the data that underpins it should be used. However, there are many healthcare professionals and innovators who are taking a thoughtful approach to ensure the delivery of ethical and trustworthy AI that has a positive impact on the lives of patients around the world. We are now building the practices, companies, and thought leaders that will shape how healthcare is delivered in the coming years, and I want to take this chance to explore some great use cases of AI in healthcare.

Source: CBINSIGHTS State Of AI Q3'21



AI FOR MEDICAL IMAGING

In recent years, deep learning (DL), in which neural networks learn patterns directly from raw data, has achieved remarkable success in image classification. Medical AI research has consequently grown across specialties that rely heavily on the interpretation of images, such as radiology, pathology, gastroenterology, and ophthalmology.

Within these specialties, AI can be used as a tool for case triage and image interpretation: the algorithms support a clinician in reviewing images and scans. It is ***reported*** that this allows radiologists, cardiologists, and other medical professionals to prioritize critical cases, avoid potential errors in reading electronic health records (EHRs), and establish more precise diagnoses. These solutions can add value both in the field with limited access to care and in traditional settings.

Deepecho is an AI medical imaging startup for ultrasound imaging that assists in medical diagnosis and preventive and predictive detection of prenatal complications. Deepecho focuses on the diagnosis of fetal growth restriction (FGR), which is one of the most important determinants of infant mortality in developing countries. The prevalence of FGR varied ***between 2.6 and 59.2% in Africa***, being one of the leading causes of newborn children in the continent. Deepecho is building software to prevent birth defects, address preterm birth, low birth weight, and their outcomes. They are simplifying ultrasound so that minimally trained clinical professionals can do the work of a trained sonographer guided by AI technologies and use it in nontraditional healthcare settings.

On the other hand, ***Dyad Medical*** develops clinical applications using artificial intelligence to enable doctors to more efficiently and effectively interpret medical image content. Their core focus is in cardiovascular (more people die as a result of cardiovascular disease than all types of cancer combined) and neurology. The solution supports treatment decision-making in the clinic and analysis in research, enabling researchers to optimize observations and conclusions.

Singaporean startup ***Us2.AI*** is automating the fight against heart disease, using AI software to simplify and democratize ultrasound of the heart, the most commonly used tool for the detection of cardiovascular risk. They are one of the few players to receive FDA clearance for the first fully automated solution measuring both 2D and Doppler cardiac ultrasound images.

AI FOR DRUG DEVELOPMENT

A ***recent study*** estimated that the median cost of bringing a new drug to market was $985 million, and the average cost was $1.3 billion. Beyond being costly, this is also a lengthy process: The ***average*** time from FDA application to approval of drugs is 12 years. And after all this money and time, ***only 5 in 5,000*** of the drugs that begin preclinical testing ever make it to human testing, and just one of these five is ever approved for human usage.

Drug research and discovery is one of the ***more recent applications for AI in healthcare***. It can be used to streamline drug discovery and drug repurposing processes, cutting both the time to market for new drugs and their costs. ***The Verge*** reports that experts have pointed to AI as a way to make it faster and cheaper to find new medications to treat various conditions. AI could help scan through databases of potential molecules to find some that best fit a particular biological target, for example, or to fine-tune proposed compounds. Hundreds of millions of dollars have been invested in companies building

AI tools over *the past two years.*

Quant Health Labs is a trial simulation platform for hematology-oncology, capable of conducting entire clinical trials in-silico, by simulating the outcomes of approved and novel treatments, patient by patient, across an entire cohort. They combine over 1 trillion data points across the clinical and pharmacological domains to tackle some of pharma's biggest tactical and strategic challenges.

Established players are making bets in the space: a new Alphabet company will use artificial intelligence methods for drug discovery, Google's parent company *announced last November.* It'll base its technology on the work done by DeepMind, another Alphabet subsidiary that has done ground-breaking work *using AI to predict the structure of proteins.* The new company, called Isomorphic Laboratories, will focus on identifying new pharmaceuticals.

AI FOR REMOTE DIAGNOSIS

Artificial intelligence in medical diagnosis can be used to support specialists with decision making, triage critical findings in medical imaging, flag abnormalities, provide specialists with the tools to prioritize life-threatening cases, and help with the management of chronic diseases. But with the rise of Telehealth and care moving outside of the hospital, it's essential to leverage the power of AI to accurately diagnose remotely.

Accurate diagnosis is of extreme importance in global healthcare systems where patients are often underserviced or overserviced. According to *some sources*, the overall rate of misdiagnosis in primary care is thought to be between 10% and 15%. Other studies report 5% of outpatients receive an incorrect diagnosis, with errors being particularly common for serious medical conditions and carrying the risk of serious patient harm.

AI and machine learning have emerged as powerful technologies for assisting with remote diagnosis, essentially helping doctors deliver more precise diagnoses or automating it all together. The future of interconnected devices and diagnoses made outside the hospital is being made possible by companies like *Hyfe AI*. Hyfe AI detects and tracks cough in real-time, remotely, and at scale: Hyfe's vision is to use smartphones and acoustic epidemiology to identify illness and track health, transforming healthcare systems for good. In partnership with the University of Navarra in Spain, they have been part of a ground-breaking project focused on digital acoustic surveillance for the early detection of respiratory disease outbreaks.

AI TO INCREASE PRODUCTIVITY AND PATIENT OUTCOMES

One of the main problems in healthcare is the lack of data interoperability, which coupled with the reliance on manual inputs of data on many different systems, creates folders and data lakes filled with inaccessible information.

Many patients have zero historic medical information when they shop for care at different providers, even if they were hospitalized 10 days ago at a competing health system two blocks away from their provider. The lack of information severely impacts patient safety and speed of care, increasing the overall cost of care delivered. Doctors are sometimes able to access data, usually by phone or fax rather

than electronically, which takes time and does not address an underlying wasted opportunity: all of these data lakes can be used to train AI and machine learning models to triage care, find billing problems or generate a rich internal database with insights that can inform decision making.

There is a great opportunity to combine these isolated sources of data with medical and governmental publications to parse risk in real-time. ***Player***s like Sorcero focus on streamlining the document intake process by extracting data from intake forms, IDs, and other insurance docs without the need for human intervention. The solution improves patient outcomes by empowering life sciences experts to dramatically increase productivity and find hidden patterns. The platform allows experts to explore vast libraries of unstructured medically relevant content, honing on meaningful analytics and takeaways.

CONCLUSION

The highlighted cases are just a small sample of the great impact and potential that AI has within healthcare. Beyond these cases, AI can be implemented to improve workflows and operations, to assist staff with automating repetitive tasks, to improve patient literacy and inform alternative care pathways, to develop innovative treatments and therapies targeted to the needs of each patient, to change the way researchers and professionals access information, and many other use cases.

The ability of AI systems to leverage vast amounts of complex, previously unaccessible ambiguous information into insight can be used to tackle some of the largest problems in healthcare worldwide. However, we should be careful in how we develop and deploy these technologies: to allow AI systems to deliver benefits across the value chain, we first need to build guidelines that help us enforce moral values, ethical principles, and avoid inherent biases in the existing data. There is a long road ahead, with much-needed research and regulation, but I am optimistic that AI will be a force for good in healthcare.

BIO:

Diego is an Investor at Plug and Play, currently scaling Plug and Play's footprint in Africa. He focuses on sourcing and leading investments across the continent, including working with the Smart Cities programs in Morocco and Egypt. Before this, Diego was based out of the Abu Dhabi and Valencia offices, focusing on Healthcare. He sourced over 300 early-stage healthcare startups globally. He has worked with leading healthcare corporations like Pfizer, Roche, BMS, and others.

Before joining Plug and Play, he was in a growth role at Wikifactory, a social platform for collaborative product development. He is an honors graduate with a BA in Interactive Media and a Minor in Art History from New York University. Born and raised in Madrid, he has lived in Abu Dhabi, Shanghai, and New York.

©Elle Magazine

IN AFGHANISTAN WOMEN ARE FIGHTING A WAR AGAINST AN INVISIBLE ENEMY - BREAST CANCER!

In the UK, an estimated 80 percent of women survive for at least five years with breast cancer, and an estimated 90 percent of women will survive in the US during that same time frame. But the disease is often a death sentence for Afghan women.

In Afghanistan, there is only one oncology department in the entire country. There is no readily available chemotherapy, and no radiation.

Here, breast cancer is the second leading cause of death of women behind maternal mortality, according to the World Health Organization and medical experts. And while maternal mortality rates have decreased significantly, the rates of breast cancer survival are not improving. Sound statistics are hard to come by in Afghanistan, but there's evidence that more women might die from breast cancer than they do from war.

To make matters worse, there is only one practicing oncologist in all of Afghanistan, according to doctors here – Dr. Zabi Stanekzai, head of cancer diagnostics at Jamhuriat Hospital.

Thirty-three-year-old Dr. Sohaila Niazi, however, plans to become the second – and the country's first-ever female oncologist. She trained in Pakistan and specializes in cancer treatment.

Dr. Niazi hopes to travel abroad soon to complete an oncology residency program, which is not offered in Afghanistan, though she says she needs a financial scholarship to do so. At Jamhuriat Hospital, Sohaila works alongside her 27-year-old sister Najia, who says she's the country's sole cognitive behavioral therapist working with patients diagnosed with cancer.

How did we get involved?

Tarry spoke with Dr. Niazi and was determined to solve this with AI. See his full video here:

To meet this challenge through the use of artificial intelligence (AI), deepkapha AI Research Lab successfully carried out a pilot project in partnership with EPOS.de a german health management group, a subsidiary of GOPA Consulting Group - which specializes in large-scale medical projects across the world.

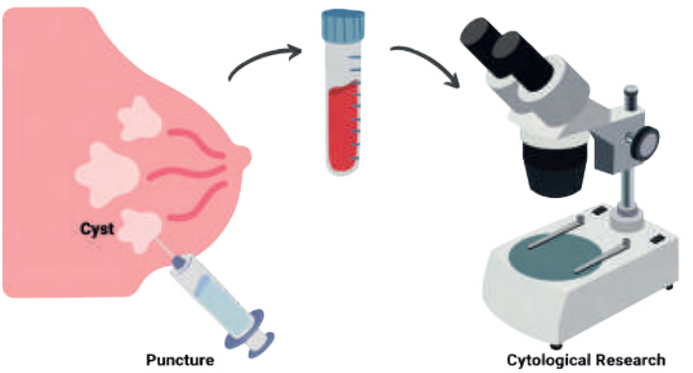
The project's goal was to use the latest AI (artificial intelligence) techniques and automatically identify malignant and benign cancer cases from a collection of cytological images of real patients in the country.

Before building a diagnostic model, it was important to understand the advantages and disadvantages of the procedure by which the model is basing its decision on. Breast cancer can be diagnosed by imaging such as an ultrasound and/or through a biopsy such as FNAC or CNB. In developed countries, the usage of FNAC for breast lesions has been reduced in favor of CNB for a variety of factors such as difficulty in diagnosis, but FNAC still has key advantages in its ease of use, inexpensive cost, and minimal infrastructure requirements (Hukkinen et al 2008).

Challenges we faced

One of the challenges of FNAC is the ability for a pathologist to make a definitive diagnosis. Accurate interpretation of FNAC results takes very extensive training and can still be somewhat inconclusive even for skilled physicians (Willems 2012). CNB results are generally easier for pathologists to diagnose and usually have better performance metrics. Although FNAC diagnosis does have very high sensitivity and specificity

(Gary et al., 2010), a meta-analysis conducted by Willems et al. showed that CNB has higher sensitivity and specificity than FNAC for identifying breast lesions. Samples which are extracted via CNB also preserve the cell structure of the lesion unlike FNAC. This allows for histological diagnosis which can give pathologists more insight into precisely which type of breast cancer the patient may have. Histopathology also allows for tumor grading, unlike cytopathology where there is not a definitive consensus of what criteria can be used to grade FNAC results (Khan 2003).



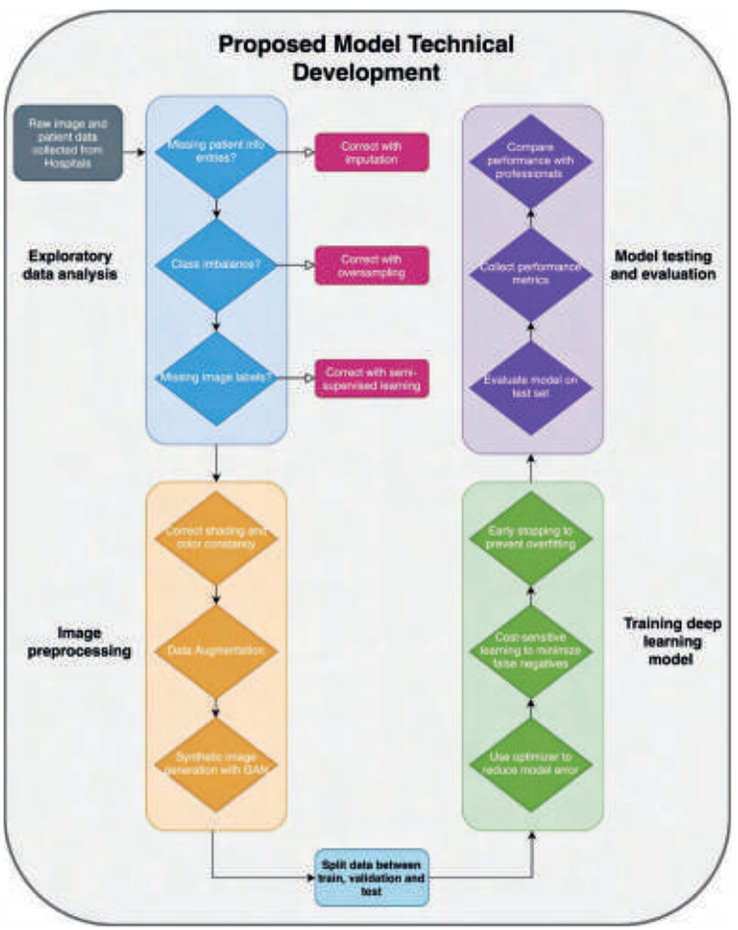


Figure 2: This is the flowchart of how the proposed model will be developed. First we will perform exploratory data analysis on the dataset to determine which techniques will be needed to ensure the model will perform well. Then, we will preprocess the images and split them between the train, validation and test sets. Lastly, the model will be trained and its performance will be compared to that of healthcare professionals.
Source: CytoDeep Feasibility Study / An EU funded deepkapha AI Research

This process applied Deep Learning (DL) on several hundred of cytology images using an advanced form of computer vision. To improve the accuracy of the AI algorithm model, deepkapha.ai applied the advanced technique of "Transfer Learning", using trained and benchmarked models to achieve an accuracy of 82% in a first step and ultimately up to 95%.

We naturally ran into a lot of challenges. First of all, there weren't many experts in cytology domain so we assembled a pool of AI researchers and engineers operate at the cross-section of technology of domain

expertise and deep learning latest techniques. We then brought together a pool of medical doctors and deep learning engineers - some of them even having double degrees as MD and Deep Learning researchers, which gave us a unique advantage of fast-tracking our access to domain experts such as cytologists and histopathologists.

Deepkapha team will now enhance, scale up and improve their joint research to create a reliable tool that doctors in developing countries can use with ease and confidence.

Author- TARRY SINGH

Tarry Singh is Chairman, CEO, and AI Researcher of AI company Real AI Inc and deepkapha.ai. Tarry has 25 years of experience working with data and has advised CxOs of global organizations to set up data-driven organizations from scratch. He speaks regularly at global AI leadership summits worldwide and conducts workshops on a regular basis with his TAs who are currently PhDs in various disciplines such as NLP, Computer Vision, Robotics, and other Artificial Intelligence disciplines.

CEO Startup Showcase Interview with AVI VEIDMAN

Co-Founder and CEO of Nucleai

BIO:

Avi is the co-founder and CEO of Nucleai, a spatial biology company helping pharma develop drugs treating cancer and improve treatment decisions.

Avi has over 25 years of experience in technology leadership roles in elite technological units in the Israeli Intelligence forces (9900, 8200). During his service, Avi established an AI-Data science department specializing in computer vision and spatial information analysis that automates satellite image analysis and led at this position 150 researchers, developers, and subject matter experts. Avi holds both an MSc. and BSc. in Geoinformatics (Cum Laude) from the Technion.

What is Nucleai, and how does it fit into the modern medical ecosystem?

Nucleai is an AI-powered spatial biology company with a mission to transform drug development and clinical treatment decisions by unlocking the power of pathology data.

The company is working with more than ten pharmaceutical companies to harness spatial biology for new drug development, clinical trials, and clinical treatment decisions. In addition, Nucleai's platform is being used to retrospectively and prospectively better stratify patient populations and improve the probability of success in clinical trials.

Nucleai is providing a complete solution that brings computational power and scale to obtaining the spatial biology-based profile of a patient using pathology data – a breakthrough that enables a new kind of framework to model biology in significantly better ways to predict patient response with higher-quality predictive biomarkers, find new targets, and develop pathology-based companion diagnostics.

Nucleai leverages proprietary, multimodal data sets of pathology images and clinical data from leading hospitals and health maintenance organizations in the U.S. and Israel to advance its platform continually. The company has a distinct advantage with direct access to more than 20 million pathology slides with tissue analysis through its AI engine. This data accessibility is being combined with Nucleai's technological innovation and the agility of its highly experienced, multi-disciplinary team to redefine how pathology can be conducted.

Tell us about the role and applications of AI in precision medicine.

AI is a powerful tool that has already shaped the world of precision medicine and will continue to do so as more advancements are made. Beyond image analysis capabilities, artificial intelligence can be employed to identify potential drug candidates, aid in detection of disease across diagnostic methods, process electronic health and clinical records, and more generally analyze, compare and integrate large amounts of varying data modalities to organize both raw data sets and processed data repositories.

Tell us about Nucleai's ATOM platform.

Nucleai's ATOM platform provides both known biomarker quantification and novel spatial biomarker discovery capabilities. The platform's AI-powered models work across multiple indications and stain types to register digital pathology images, perform QC analysis, segment tumor and stromal areas of interest, and identify and quantify cell types and biomarkers. ATOM then leverages its computational power to calculate spatial features, like distances between cells of interest, immune cell densities, and areas of coexpression, to name a few. Finally, integrating clinical, outcomes, and omics data allow ATOM to identify predictive spatial biomarkers that impact outcomes and treatment decisions.

How is Nucleai bringing the power of spatial biology to reality? How can pharmaceutical companies and clinicians use AI-powered image analysis applications?

Nucleai's diverse team of biology and data science experts work in tandem to create the AI-powered models that drive forward our mission of unlocking the power of spatial biology. Our technology goes beyond standard image analysis and biomarker quantification to reveal complex pathological features through automated structuring, identification, and feature calculations. The power of spatial biology lies in the correlation and integration of those features with clinical, genomics, and outcomes data. This is where novel biomarkers and impactful insights can be uncovered, strengthening drug development and treatment decisions.

Pharmaceutical companies, biotechs, and clinics can use Nucleai's AI-powered image analysis to provide the computational power needed to unlock the valuable layers of spatial information to inform drug development, laboratory practices, clinical trial enrollment, and treatment decision making.

Can you shed some light on Nucleiai's offerings, including its exploratory research platform and AI-powered clinical trial assays?

Nucleai's ATOM platform uses cases across research, clinical trials, and diagnostics. Starting in exploratory research, ATOM can be deployed to quantify known biomarkers, better understand mechanisms of action, and identify novel spatial features of interest for further analysis and research. The insight gained at this step equips research teams with an arsenal of spatial data and guides the next steps of their drug discovery efforts. Moving into clinical trials, ATOM can be used to analyze data on closed trials, correlating pathological status and discovering novel features that may predict response rates or other clinical endpoints. For prospective trials, our offering can be used first to gain an understanding of what types of patients are best suited for a particular treatment and then screen patients to identify the specific population for enrollment into that trial. Our last offering centers around companion diagnostic development, which provides a tool in the clinical setting to efficiently determine the patients that will respond to a particular therapeutic given their unique tumor spatial signature identified by ATOM.

What are some of the challenges and trends you foresee in precision medicine and AI?

One large challenge that we face in precision medicine and especially AI-based image analysis is digitizing data, especially Whole Slide Images. Our imaging technology relies on tissue samples to be viewed and imaged digitally, which is not yet a universal practice for pathological workflows. Although it is becoming more and more common to image, view, and store pathological samples digitally, the slow adoption can create a roadblock for research teams to work with companies like Nucleai. Looking ahead, the industry is trending towards full digital adoption of all data modalities, which will enable AI imaging companies especially to more easily support the pathological components of drug development.

Tell us about Nucleai's latest partnerships in Healthcare and AI.

Nucleai has a robust set of data partners across the healthcare ecosystem with proprietary access to pathology images and clinical data. Israeli based partners like Sheba Medical Center, the most extensive hospital system in Israel, and Clalit, one of Israel's largest health care providers, as well as American cancer centers, such as Fox Chase Cancer Center and Jefferson Health, allow Nucleai to tap into real-world data that we leverage to build and train our models. The sheer amount and variety of data made available to Nucleai through these partnerships allow us to maintain a suite of pre-trained algorithms that can be deployed for our customers right away. In addition, the heterogeneity of our training data sets equips our models to be successful regardless of the study type or available data from the customer.

LOOKING FOR BYTE SIZE AI NEWS ?

CHECK OUT

MARKTECHPOST.COM

Marktechpost is an AI News Platform providing easy-to-consume, byte size updates in machine learning, deep learning, and data science research. Our vision is to showcase the hottest research trends in AI from around the world using our innovative method of search and discovery

Exclusive Interview with KIMBERLY POWELL VP and General Manager of NVIDIA Healthcare

Kimberly Powell, Vice President of Healthcare

Kimberly Powell is vice president of healthcare at NVIDIA. She is responsible for the company's worldwide healthcare business, including hardware and software platforms for accelerated computing, AI, and visualization that power the ecosystem of medical imaging, life sciences, drug discovery, and healthcare analytics.

Previously, Powell led the company's higher education and research business, along with strategic evangelism programs, NVIDIA AI Labs, and the NVIDIA Inception program with over 8,500 AI startup members.

Powell joined NVIDIA in 2008 with responsibility for establishing NVIDIA GPUs as the accelerator platform for medical imaging instruments.

1. Tell us about your journey at NVIDIA. How has the advent of the NVIDIA GPU transformed the application of AI in healthcare?

My journey at NVIDIA started 14 years ago in the medical devices sector. When I started, NVIDIA was primarily known for computer graphics, and over time, NVIDIA has expanded into other areas, including supercomputing and artificial intelligence.

GPU & Computer Graphics

NVIDIA's foundational invention was the graphics processing unit ***(GPU)***. The GPU's purpose is a very high-level, parallel processing unit to run certain applications at orders of magnitude faster than CPUs or other architectures. The GPU is what really got me excited about joining the company when I did about 14 years ago. This type of invention creates paradigm shifts in industries.

The first killer application of GPUs was computer graphics. In fact, our first application in healthcare was for computer graphics and radiology. Radiology is a field where we use devices to see inside the human body. We wanted to be able to see things in more and more detail, with advanced imaging like in 3D MRI.

Accelerated Computing & Supercomputers

About fifteen years ago, NVIDIA expanded beyond computer graphics into an accelerated computing company. GPU acceleration was paramount for the world's supercomputers. Supercomputing is an area that we are still heavily involved in today. NVIDIA is powering over 70% of supercomputers, which is pretty incredible.

One of the most important application areas of supercomputing centers globally is Life Sciences. One of the greatest challenges of humanity is to understand disease. At NVIDIA, we do that through very large-scale bioinformatics, molecular modeling, and simulation. This is the tip of the spear of what you could imagine industries like the pharmaceutical industry taking on. We always engage at that whole ecosystem level, starting at research, so that we can be at the bleeding edge of what our industry is going to look like in 5 to 10 years.

Domain-Specific Artificial Intelligence

Now, in 2022, the biggest and fastest-growing application area is artificial intelligence. AI is going beyond graphics in terms of what it's doing for our company. It is the biggest technology force of the current time. We have always firmly believed that. Now the mission statement of NVIDIA Healthcare is to bring that capability of artificial intelligence to the healthcare industry.

If you think about AI and the notion of intelligence, it means that it's domain-specific. There is a reason why doctors go to school and practice for decades before they are considered a specialist; because it's very domain-specific. That is what we are doing in the healthcare industry. We are taking these computational approaches that NVIDIA has pioneered from computer graphics to accelerated computing and artificial intelligence and putting them in the hands of the healthcare industry.

Back 14 years ago, when I started the healthcare practice for NVIDIA, we were getting all these early indicators saying that this sensor technology that was being invented needed a step function in terms of its computing power. All these improvements in the sensor technology put this huge strain on the downstream processing and the human interpretation of all of that data. Today, you will ***see NVIDIA inside of all of your modern medical devices, including CT, MRI, ultrasound, genomic sequencers, and microscopes.***

Artificial intelligence is becoming the computational workhorse for medical device innovation. It is an area that NVIDIA is really, really focused on and we have built computing platforms to support this.

2. How would you describe NVIDIA's role in contributing to the ecosystem of medical imaging, life sciences, drug discovery, and healthcare analytics?

Healthcare is a major industry. Those are four segments of the healthcare ecosystem that are also giant in nature. NVIDIA started in medical devices and it is still one of our core areas of contribution. The nature of medical device innovation has all of this high throughput data, and this is what is triggering the digital biology revolution. Genomics is one of the most intense data science areas ever because of these 3 billion letters that make up the story of each individual human.

NVIDIA has this unique view of being a computing platform company. We think about medical devices and allowing them to become more sophisticated in what they can sense and what they can build into their sensors. We want to help the medical device sector innovate. By helping the healthcare industry in creating all of this downstream data to really understand human disease, and then applying it to the challenges in each one of these industries.

Drug Discovery & Genomics

In drug discovery, you think about what steps need to be taken. We first have to identify the target, then identify what molecules might affect the behavior of that target protein, then tie that all together in ways that are completely in silico (completely done in the computer). We use artificial intelligence, modeling, and simulation so we can reduce the amount of expensive time-consuming, error-prone experimentation that has been been used previously in early-stage drug development. You look at genomics being at the front of that pipeline, to really teach us about and help identify the genes that code for the proteins that cause our body to do certain things - good and bad. And then all the way through to these very, very large simulation problems. If you look at the drug discovery process that encapsulates it all: We go through genomics and proteins and molecular simulations, all the way through to clinical trials. NVIDIA is instrumental throughout the process- from identifying the genes and variants, all the way to early-stage in silico drug development, finishing with clinical trials.

Doing Things In Silico

How is this manifesting inside of an actual human, studying that and bringing it all back again and creating this sort of loop? At NVIDIA we are trying to, as much as we can, put the ecosystem and processes in silico. Computer science approaches – whether it be scientific computing, artificial intelligence, or advanced visualization technique – can be applied to this data in new and sophisticated ways. This data processing is well beyond what a human, any one human, could really endure, take on, and make decisions about. So doing things in silico is really how we think about it.

3. Can you tell us a little about NVIDIA Clara and other healthcare tools that NVIDIA is working on? How are these tools impacting elements of the healthcare industry, such as radiology, medical imaging, and genomics?

NVIDIA is a whole stack computing company. This has really helped the pharmaceutical industry understand us a bit more. Some people just call us a chip company, and we obviously find that really flawed with all the work that we are doing now.

Layer 1: GPU, chips, systems, data centers

The first layer of the stack is really around the GPU. However, we have also moved well into full-on systems. That first layer is about chips, systems, and whole data centers as our product.

Layer 2: Acceleration Layer

The second layer is the acceleration layer and is where the acceleration comes in. How do you take advantage of that architecture at more than a chip-level? At multiple chip levels? At multiple node levels? That acceleration layer is what really put NVIDIA on the map in the area of artificial intelligence, being able to do this deep learning at very, very large scale. A lot of the things we build at this second layer can be used in financial services, in autonomous vehicles, in the omniverse.

Layer 3: Industry Application Framework Layer

In these last 5 to 10 years we are developing the third layer of our company, and we call that the industry application framework layer. The goal of the third layer is to take that acceleration and system layer, and make it more domain-specific.

For the healthcare domain, we call that industry application framework ***NVIDIA Clara***. This framework was named after Clara Barton, the inventor of the American Red Cross, and we think of Clara as a platform to help people. Our Clara platform builds upon those first two layers below it – the GPU system layer and data center layer – and leverages everything we do, as a 20,000 plus company. In this third layer of Clara, we focus on a few very specific areas.

Clara Holoscan:

Medical devices are our core. Over the last several years, we have developed our Software-Defined computational platform for medical devices called ***Clara Holoscan***. This is where we are building actual specific systems so medical devices can do the end-to-end workloads that they need: everything from very high throughput sensor processing, to all of the AI processing they want to do in-device, and even doing visualization. Think about an ultrasound machine: You have a sensor in and display out all on the same machine, and all of the AI and image processing that has to happen in between. This is a very typical pipeline for medical devices.

We are building a computational platform so that medical device manufacturers will not have to think about the nuts and bolts in this. It really builds upon NVIDIA's three core engines that we have now. (1) We invented the GPU. (2) We have been pioneering ARM-based processors. Our ARM CPU architecture is what powers all of our self-driving but now can be used for things like medical devices. (3) The third one we recently added to our family with the Mellanox acquisition is our data processing unit (DPU). You need to get data into the node at very, very high speeds. So we now have this three-engine architecture. And that's, again, a very unique position of NVIDIA.

We want to make it much easier for the medical device community to take advantage of those three engines, and to really help them accelerate their innovation in that space. So Clara Holoscan is exactly that. The system architecture to the acceleration layer that sits on top of that to the domain-specific applications. If you look at Clara Holoscan, we have reference applications for endoscopy which is what is powering this minimally invasive surgical market. Clara Holoscan is one of the places that we are really, really excited about, and you are going to see a lot of upcoming development in that area.

We want these medical devices to become a self-driving car in a sense. What do I mean by that? We want to move into the software-as-a-service business model. Companies do not want to sell an instrument once and have to maintain it for 10 years. They want to be able to continue to innovate on AI applications and increase value upon the instru

ment that goes in it. That computational platform we built allows for that software-defined, medical device era to come to this field. Much of what we have learned from self-driving is completely applicable to the medical devices market.

NVIDIA Clara Discovery

Another area that we are greatly focused on is taking all of the work that we do in the supercomputing industry, and everything that we have learned from artificial intelligence, and bringing it into the drug discovery market. In the last 18 months, we announced ***NVIDIA Clara Discovery***, our computational platform for drug discovery. Clara Discovery is all about the bleeding edge of AI and applying it to this very unique data within medical devices and biotech. We are looking at data such as protein sequence data or smile strings that represent a molecule in the bioinformatics space and genomics data.

Transformers & Generative Models

I saw that Marktechpost followed transformers and generative models in 2021 and how applicable they are to these incredibly challenging datasets. Alphafold 2, enabled by transformer AI, allows you to essentially feed whole databases of protein sequence information so you can predict the structure of a protein.

We are pioneering generative models for molecule generation with AstraZeneca using something we call ***MegaMolBART***. Using transformers in generative models to go beyond the molecular databases that exist, because there are 10^60 potential molecules that we could build. Our databases are still quite small and we want to be able to explore as much as we can. It has a lot of downstream applications in the drug discovery space.

MONAI & Computer Vision

I think most people know NVIDIA in the healthcare space on the imaging side. What we have done over the last four years is build an AI framework for medical AI. We call this framework the Medical Open Network for Artificial Intelligence, or ***MONAI***. MONAI is a PyTorch-based framework for deep learning in healthcare imaging. MONAI is largely targeting a lot of the imaging applications – such as radiology, pathology, or real-time video – used in the surgical space. This week (January 19, 2022), we surpassed 200,000 downloads of this framework. It has all the domain-specific data ingestion, transformation, and model architectures used in this space.

How do you deploy this into a clinical environment so you can validate it? We are working with a huge consortium of contributors with MONAI and building this application framework because we want to make it very, very accessible.

Computer Vision was one of the frontrunners of the application of AI. Computer vision applications in healthcare have tremendous opportunities. Tens of thousands of algorithms will be developed to serve the radiology industry. You can use algorithms everywhere from capturing the right image to de-noising that image to then looking for anatomical structures in that image. Doing all of the things that are repetitive that humans have to do and then presenting that information such that we can help the overstretched radiologists. With MONAI, we are really excited about what's happening in that space and we continue to put a lot of focus on that.

Data access & federated learning

Secondary to MONAI and just building the applications, we are also addressing one of the main challenges in healthcare around data access. A lot of computer vision and CNN (convolutional neural networks) approaches require a lot of data, a lot of labeled data. MONAI helps with that. Healthcare data changes all the time. We want to enable a world that can better adapt to that. All of a sudden, we are seeing lungs that have COVID pneumonia that we have never seen before. How can we create robust algorithms in real-time for that? We are doing that through a federated learning platform.

NVIDIA FLARE

We recently announced open-source ***NVIDIA FLARE*** (Federated Learning Application Runtime Environment), which is our federated learning framework. We worked with a consortium of 20 different hospitals and a model that was developed at Mass General. We delivered that model to 20 different hospitals so they could contribute learnings from their data but not have to contribute any data at all. It created this really amazing multi-role model that predicts the oxygen need of a patient who had an x-ray and had some lab work done. It shows that the future of AI development will be in a federated manner. Federated learning allows you to learn from data that is happening out on the edge but to not have to share that data that's happening on the edge.

So that in a nutshell is a lot of what Clara is focused on. We also have lots of efforts in NLP but, there is probably more than we can touch on.

4. How is NVIDIA planning to use federated learning within its healthcare division?

With open-source ***NVIDIA FLARE*** (Federated Learning Application Runtime Environment), we work with a lot of collaborators. At Mass General, they had this really neat model that used two different modality types: (1) electronic health record data and (2) X-ray data. The two different modality types combined to make this prediction. After we did this program, we actually package all of the training tools and the model itself into our ***NGC***, which is essentially our AI software hub. We publish the model so that the world can take and build upon it. It is not an FDA-approved algorithm, but it is meant to be a tool to help the world build upon it: whether they want to learn how to do federated learning, or whether they actually want to take that model and build it into their own application framework to go through the FDA validation process. We see this as absolutely the future of model co-development.

NVIDIA is getting approached by a lot of the industry to do that co-development. This is a very safe way to respect the privacy of data, but move the field forward and develop cutting-edge algorithms that can be heavily used. We are also enabling all of our other industries, whether it be our self-driving car industry, our financial services industry, or our retail industry. All industries have data governance challenges. Data cannot be static. Data isn't static. If you want AI to be able to deal with non-static data, it has to learn from non-static data. We believe federated learning is absolutely going to be what call AI 2.0. Federated learning will allow us to be able to take advantage of all the data that the world is going to continuously produce in a safe way.

5. Tell us about some of NVIDIA's latest partnerships in healthcare and AI.

Cambridge-1

We have many partnerships going on. One of the ones that we are super excited about is built in the UK, called ***Cambridge-1***. Cambridge-1 is dedicated to large-scale AI research in healthcare and the first NVIDIA built for external access. One of our collaborators, Kings College London, has developed some brain disease algorithms on it and they are actually deploying it into their clinical environment. We are also working with start-ups like Peptone, and others on Cambridge-1, which is the most powerful AI supercomputer in the UK.

We are also working with AstraZeneca on these transformer-based generative models for molecule generation, which we call ***MegaMolBART.***

Genomics & Oxford Nanopore Technologies

We are working with the genomics sequencing company, ***Oxford Nanopore Technologies***. Stanford University's Dr. Euan Ashley, had a dream of being able to more rapidly diagnose critical care patients through the use of genomic sequencing. Oxford Nanopore Technologies and NVIDIA have been working together for many years. By accelerating the whole pipeline – everything from the base calling on the sequencer all the way through to the variant calling that can decide which genetic disease you may be suffering from – we can more effectively intervene with treatment. We were able to take the world record from 14 hours down to seven and a half hours. NVIDIA is working on the accuracy and the speed of genomics with Oxford Nanopore.

Global Cancer Research

Some of the other partnerships from GTC Fall include cancer research. There is amazing work that the global cancer centers are now doing with AI. There are so many unmet needs in diseases and cancer is a big one of them. We are working with MD Anderson, St. Jude, Memorial Sloan Kettering, and German Cancer Research Center DKFZ.

Working with the startup community

In the last two years, some $40 billion in funding have flown into the drug-discovery startup community, and for good reason. The breakthroughs of alpha folds and protein structure prediction, the advancement of genomics, the fact that we can do more with AI and natural language processing. It's a perfectly ripe time for these new companies to be established. JP Morgan Health just finished up. There's just partnership upon partnership of large pharma, partnering with these AI platform companies to really look for advances and acceleration.

6. What do you foresee as the biggest challenges in 2022 onward for the AI and healthcare domain and where do you see NVIDIA fitting in?

A couple of challenges that come to mind revolve around innovation, complexity, ease of use, accessibility, and the development of specific tools that address some of the challenging data problems in healthcare. Can Clara Holoscan address the innovation problems? Can all of Clara address the full-stack computing that is complex, but make it easy to use? Can we make it more accessible and state-of-the-art? Can we build specific tools and platforms that address some of the challenging data problems in healthcare?

(1) Reducing the complexity for the healthcare industry as computing complexity continues to grow. In the healthcare sector, mastering the industry involves understanding the clinical problems, the workflows of the doctors, and the patients that they are trying to serve. It is very hard to do that extremely well and stay at the bleeding edge of computing approaches. With Clara Holoscan, we are working to take that complexity and make it very easy for healthcare industry professionals to remain focused on the problems. We want to partner with them on making new computing approaches accessible to them so that:

. Their innovation can be accelerated.
. Their go-to-market can be much easier.
. They can stay innovative by moving into the software-defined, software-as-a-service business model that they so desperately need and want.

How do we reduce the complexity for the healthcare industry so that they can bring these innovations to market sooner? There are all of these AI algorithms, but they have not been productized. There are several reasons for that. We are going to make sure that a computing platform isn't the reason. You can build this application through our ubiquitous platform. You can deploy it in an instrument. You can deploy it in a data center of the hospital. You can deploy it on any cloud. NVIDIA technology is homogeneous, and you can deploy it where your business model cares to have it.

(2) Being able to stay state of the art and making it easy to do that. NVIDIA wants to make AI accessible for research and discovery. Over in Germany, we are partnered with their cancer center DKFZ so that we can give their clinician / data scientists all the tools to ask as many questions as they wish and build all the AI application models using state-of-the-art approaches.

With MONAI, we are helping doctors use AI to label images to really cut down their time of being the expert, by labeling highly required data. Our computing platforms enable that.

(3) Data accessibility problem
Third is the data accessibility problem. There are several ways you can skin that.

- Federated learning is absolutely one way we can do it. Federated learning is going to be a framework that is going to connect living breathing data to the evolution of models going forward. In the future, federated learning will enable us to develop robust models without sharing data.

- The other is state-of-the-art approaches. What's so novel about these transformers is you do not do it in a semi-supervised way (you do not have to have labeled data). For healthcare, that is huge because we will never have enough labeled data. We did have some breakthrough research ourselves at NVIDIA. You can use transformers for natural language processing, you can use transformers for pre-construction prediction. We want to use state-of-the-art approaches that help us overcome some of the data challenges.

Samsung SDS: Revolutionizing Healthcare Industry with AI

The rapid development of artificial intelligence has been a hot topic in recent years. The technology is not only invading each industry, but it's also changing how we interact with our cars and computers to help us through day-to-day life tasks like diagnosing patients or driving cars without drivers.

AI in healthcare is also a powerful tool that can be used to help doctors and patients. Instead of relying on trial-and-error, AI systems are able to analyze vast amounts of data with pinpoint accuracy at speeds comparable to or faster than humans.

Medical diagnosis using machine learning and other cognitive disciplines is an important AI use case. With patient data, AI can help doctors deliver more accurate diagnoses while also making proactive recommendations for preventive care to patients.

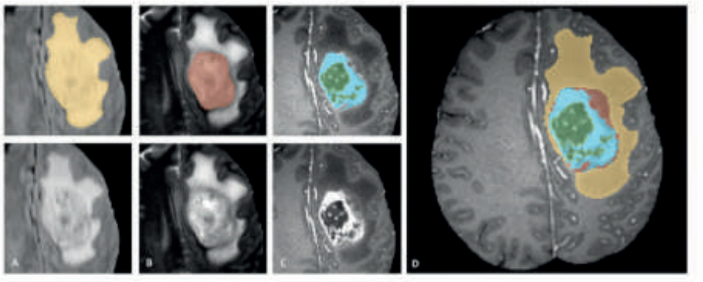
The race to create AI-powered healthcare solutions is on. Most of the big tech unicorns are involved in AI in healthcare. We are featuring one of the cool stories of AI in Healthcare from Samsung SDS.

Samsung SDS: A Global Leader in Software Solutions & IT Services

Samsung SDS America (SDSA) is the U.S. subsidiary of Samsung SDS, a global software solutions and IT services company that is a global leader in digital transformation, AI transformation, and innovation solutions. Samsung SDSA helps organizations optimize their productivity and make smarter business decisions using enterprise software solutions for secure mobility, HPC managed services, DOOH, and advanced analytics and contextual marketing. SDSA enables customers in government, financial services, retail, and other industries to drive business and improve their competitive positions in a hyper-connected economy with differentiated industry-specific enterprise solutions focusing on the IT technology of analytics, IoT, mobility, and blockchain.

AI Healthcare Research at Samsung SDS

SDSA have completed several AI research efforts and proofs-of-concept focused on classifying, detecting, and segmenting pathologies/diseases/cancer in images and videos. In doing so, SDSA have also developed state of art labeling models in 2D and 3D, and this research applies to all areas of medical imaging from MRI and CT scans, over X-ray, ultrasound, and microscopic images to regular photos.



AutoLabel: Automation of tasks for healthcare professionals using AutoLabel.

Before anyone can develop an AI model, a training dataset needs to be labeled. SDSA's AutoLabel Software-as-a-Services (SaaS) applies an unsupervised process before Active Learning in order to speed the creation of labeled datasets by selecting the most rele-

vant sample for humans-in-the-loop to label first. The result is a solution that requires only a fraction of the sample set (10%) to be labeled by humans, while the rest is auto-labeled. AutoLabel either requires a smaller set of data for the same accuracy or achieves higher accuracy for a given set of data. The result is a lower Total Cost of Ownership (TCO), significantly shorter time-to-train, and a more accurately labeled dataset which leads to more accurate AI models.

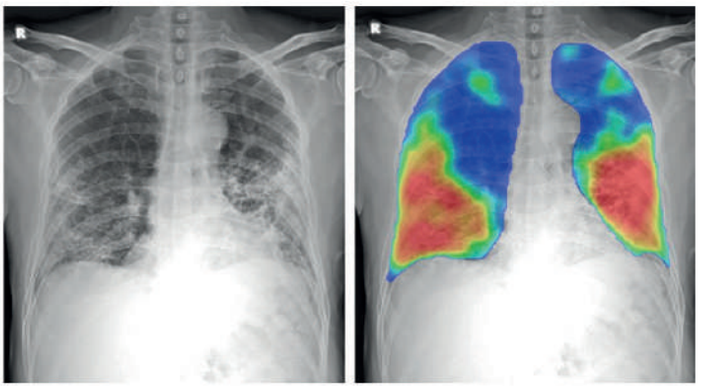
A big focus for Samsung SDS has been the healthcare sector, where AI has been automating imaging processes in areas such as radiology. As Patrick Bangert, VP of AI, states in no uncertain terms: "The medical system worldwide is overloaded. There are not enough doctors around to fulfill the demand, and they are heavily impacted by workflows, bureaucracies, and billing processes. The amount of time that any physician can spend with a real human patient is really a minority of their time. Also, a doctor's opinion might be wrong about 30% of the time, and so complex diseases – diagnosed from complex medical images – are often either missed or mis-diagnosed. Getting a second opinion is already a billion-dollar market today."

Second opinions typically come from other humans. Artificial Intelligence provides an objective third party and very neutral second or third opinion to any diagnostic procedure. Using AI removes the variability of any one physician having a certain amount of experience, because the AI model can process vastly more cases and does not forget things.

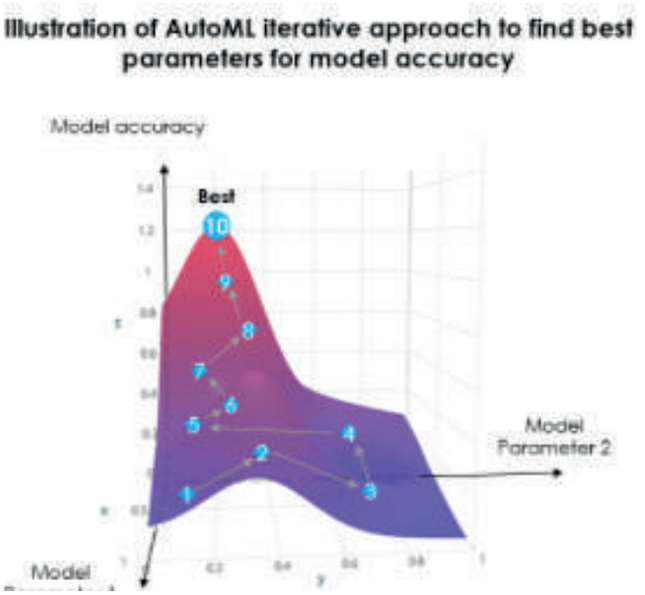
Ricky Datta, the director of AI engineering at Samsung SDS, observes; "AI models typically achieve an accuracy in the high 90% range, sometimes up to 99.9% accuracy in certain cases, depending on how much data is available and how good the quality that data is. By automating these processes, we simply make everything faster. The benefit of this would significantly lower the costs from both providers and patient's care costs, while increasing accuracy of diagnosis and speeding up the overall time from the initial pain point to the start of the treatment cycle.

Detecting COVID-19 in X-ray images using SDSA's AutoLabel solution

Recently, SDSA's AutoLabel solution was used to detect COVID-19 disease in X-ray images of lungs in order to reduce the number of samples required to be labeled to achieve an accuracy asymptote. Only 6% of sample images were required to be labeled in order to produce a model with high accuracy which lead to a 94% savings in labeling effort versus typical methodologies. In addition the resulting SDSA model is 4% more accurate than the RAPID Nose Test which takes 15 minutes for diagnosis, and 1% more accurate than the PCR Test which takes 1-3 days to return results. The SDSA COVID-19 solution requires only as much time as it takes to take an X-ray and provides more accurate results.



Automated Machine Learning with Brightics AI Accelerator



Once a dataset is labeled, scientists may train and optimize a model on that dataset in order to perform its specific use case, accurately. Brightics AI Accelerator is a subset of Samsung SDS that provides the simplest, fastest, and easiest Automated Machine Learning (AutoML) platform for professional Machine Learning (ML) and Deep Learning (DL) scientists who want to reduce model training from weeks to hours with only a few lines of code. Progressive enterprise companies are turning to automated machine learning to increase productivity efficiently with each additional server in a distributed cluster without having to build and maintain infrastructure.

AutoML automates and accelerates model training on tabular data by using automated model selection from Scikit-learn, automated feature synthesis, and hyper-parameter search optimization. AutoML with synthetic feature generation exploits up to 512 CPU cores simultaneously to produce a Scikit-learn model in 1 hour versus 2 months using traditional methods.

AutoDL automates and accelerates deep learning model training using data-parallel, distributed synchronous Horovod Ring-All-Reduce TensorFlow and PyTorch frameworks with minimal code. AutoDL exploits up to 512 GPUs per iteration to produce a model in 1 hour versus 3 weeks using traditional methods. AutoDL automates transfer learning for image data considering all models in the model zoo with hyper-parameter search.Distributed training can compress model training ~13x by using 8x more GPUs for NLP use cases, and, based on a model template, AutoDL/AutoML will help Data/M-L/DL Scientists identify relevant parameters, automatically.

SDSA's Dir of AI Sciences, Hankyu Moon states, "The continued R&D investment in AI generates a great amount of AI model and algorithm choices toward solving real-world problems, including healthcare use cases. These AI solutions are ever-expanding and growing in technical complexity. Due to the successful application to some of the medical use cases, the expectation also grows about further adoption to a wider scope of healthcare use cases. It takes deep expertise both in AI and healthcare plus compute/time resources to identify feasible solutions from these model choices and empirically validate and reduce clinical risks." Moon further explained that another key part that AI can play in healthcare imaging is hyper-parameter tuning: "The algorithm that trains one of these models has parameters of its own. They are set typically by a human being, which then leads to a trial-and-error process of tuning these things, correctly. AutoML/DL tunes them for you, automatically."

Hyper-parameter Optimization Benefits

- Saves 80% of Data Scientist time in parameter exploration
- Data Scientists can manage 4x more simultaneous projects; while refining 1 model, automate parameter optimization for 4 other models

Approach:

- Automated, brute force exploration of all combinations
- Finds good parameters in a systematic way
- Might not be optimal parameter but scientists can still use the parameter as a base for refinement

While AutoML consolidates a month of work into 10 minutes, AutoDL consolidates 3 weeks of work into 1 hour. Automatic provisioning and orchestration of jobs and resource clean-ups when completed saves you money.

When it comes down to the AI accelerator, distributed training is another tactic of Samsung's as Datta explains: "We utilise more than one graphics processing unit (GPU). To help speed up the process, we may use many hundreds of these GPU processors, simultaneously distributed over many computers to execute a single training task." With the AI accelerator as the go-to toolkit for Samsung SDS, the aim is to work on all the areas of efficiency, labeling, tuning of the models and distributed training in parallel, so they can "execute the artificial intelligence workflow a lot more quickly and efficiently and actually arrive at a much more accurate model in the end." said Datta.

Samsung SDSA also offers premium AI assessment and consulting services just like they have done for many other Samsung businesses over the last 10 years. Turnkey AI project delivery from initial assessment to actionable intelligence solution deployment for your specific use case and desired results is available at your request.

There is no question about the value that deep learning brings to auto-labeling pathology images. We are really excited to explore the idea of a system that can learn little by little from pathologists. A doctor takes 20 to 30 minutes to draw their annotation on top of a single image. And to train an AI algorithm to recognize anything meaningful, thousands of such annotated images are needed. The process is the most time-consuming and expensive of the entire AI workflow. It's the primary obstacle the AutoLabel facility of the Brightics AI Accelerator is there to overcome, using a technique called active learning to speed that process up by 90%, so it will reduce the amount of human labor from 100% down to 10% - a very significant gain, usually in the double-digit millions per model.

Please, visit
https://www.samsungsds.com/us/autolabel/autolabel.html or
contact us directly at
info.aiaccelerator@samsung.com to learn more, today!

From data annotation to model training & optimization to deploying and running a model at inference in production, Samsung SDSA's MLOps pipeline and services is Revolutionizing the Healthcare Industry with AI!

Meet the Key AI Team Members at Samsung SDS

Patrick Bangert, VP of AI:

"The value of AI is proven. The upcoming challenges will be encountered deploying AI at scale in the cloud, on-premise, and on edge networks and devices in the hands of users and embedded into sensors. In addition, the continued optimization of these models out in the wild will require an efficient and automated MLOps pipeline." Patrick heads the AI Division at Samsung SDSA. He is responsible for Brightics AI Accelerator, a distributed ML training and automated ML product, and AutoLabel, an automatic image data annotation and modeling tool primarily targeted at the medical imaging community. Among his other responsibilities is to act as a visionary for the future of AI at Samsung. Before joining Samsung, Patrick spent 15 years as CEO at Algorithmica technologies, a machine learning software company serving the chemicals and oil and gas industries. Prior to that, he was an assistant professor of applied mathematics at Jacobs University in Germany, as well as a researcher at Los Alamos National Laboratory and NASA's Jet Propulsion Laboratory. Patrick obtained his machine learning Ph.D. in mathematics and his Masters in theoretical physics from University College London.

Hankyu Moon, Dir of AI Sciences:

Dr. Hankyu Moon has been a Director of AI Science as part of the AI Team, Samsung SDS Research America since 2015. He received a B.S. degree in Mathematics Education from Seoul National University and a Ph.D. degree in Electrical and Computer Engineering from the University of Maryland College Park. His R&D career in AI started when he joined the Center for Automation Research, the University of Maryland in 1996 as a graduate student. Prior to joining Samsung, he worked at NEC Research Institute and Hughes Research Laboratories as a research scientist. Throughout his career he has collaborated with world-class researchers, producing numerous publications and patents in the areas of Computer Vision, Deep Learning, and Data Mining. At Samsung, he and his colleagues have developed Data-Efficient AI for several years, resulting in the successful launching of the AutoLabel solution that leverages Representation Learning, Active Learning, and Unsupervised Learning.

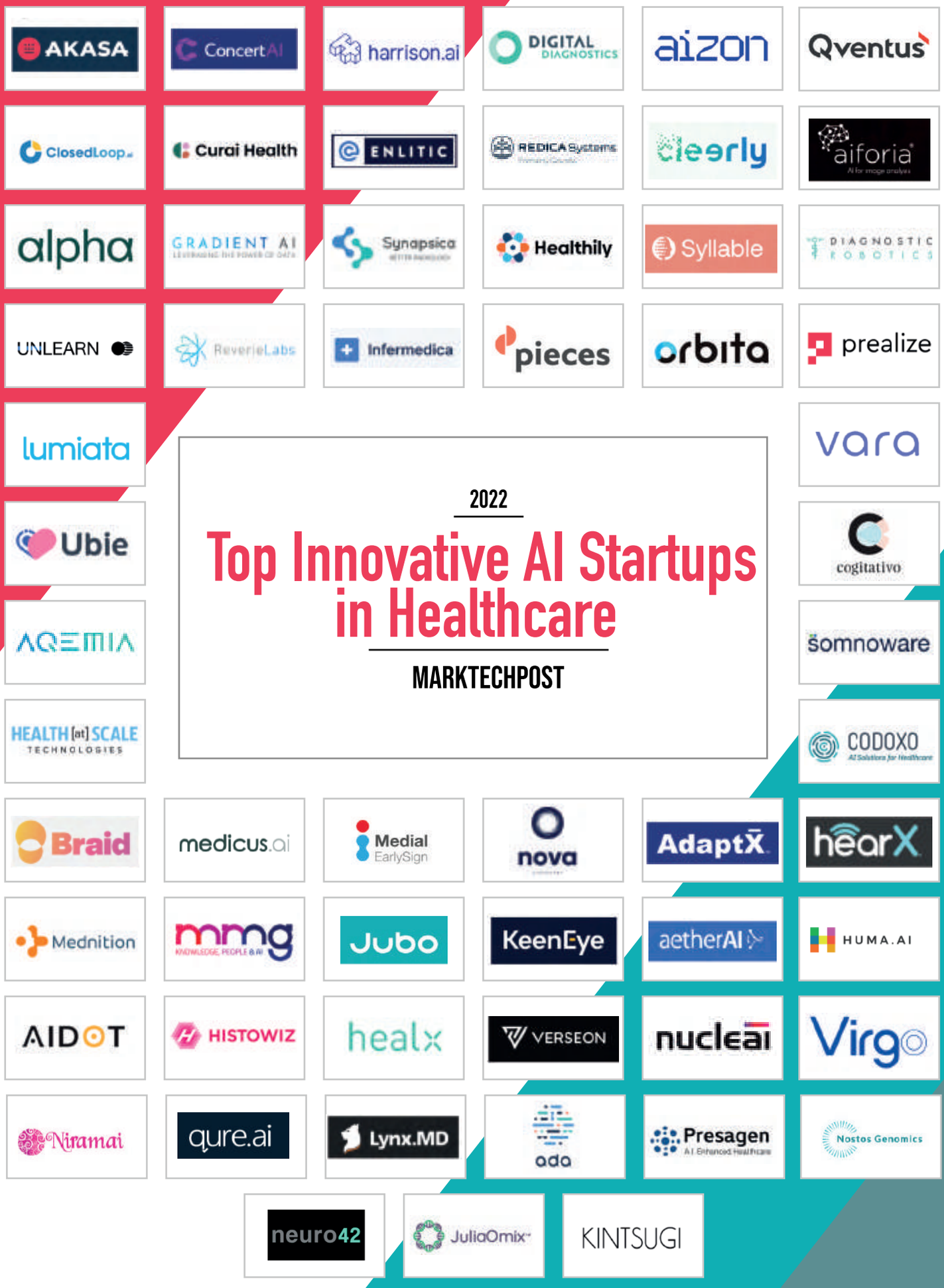
Meet the Key AI Team Members at Samsung SDS

Ricky Datta, Dir of AI Engineering:

Ricky has been responsible for the technical direction of Brightics AI Accelerator since day one. He oversees all aspects of the design and implementation of AI Accelerator. A seasoned Enterprise Software Leader, Ricky is an expert in Machine Learning, Deep Learning, and Big Data and works together with customers to ensure all their specialist technical needs are met. He started his career in small software startups to specialize in Enterprise Software. Ricky has a bachelor's degree in Computer Science from the Indian Institute of Technology, New Delhi.

Dan Waters, Dir of AI Business Development:

"The biggest challenge will be wrapping full, compelling-end-end solutions for medical users around the proven AI technology. Designing the human-computer interface in an elegant and performant manner that is a step function far and above what is available to physicians, radiologists and pathologists, today, will be required." Dan leads Business Development for the AI Team at Samsung SDS America and is responsible for the go-to-market strategy. As a Senior Business Development leader, Dan's devoted his career to bringing to market disruptive products, building teams to scale rapidly and quickly growing revenue for companies like xnor.ai, Apple, Siri, and Motorola. In 2009, Dan introduced Siri to Apple and was retained by Apple following its acquisition of Siri in 2010. Dan is an MBA graduate of the Kellogg Graduate School of Business, has a BSEE from Iowa State University, and is conversant in Spanish.


AKASA
ConcertAI
harrison.ai
DIGITAL DIAGNOSTICS
aizon
Qventus
ClosedLoop
Curai Health
ENLITIC
REDICA Systems
clearly
aiforia
alpha
GRADIENT AI
Synapsica
Healthily
Syllable
DIAGNOSTIC ROBOTICS
UNLEARN
ReverieLabs
Infermedica
pieces
orbita
prealize
lumiata
vara
2022
Top Innovative AI Startups in Healthcare
MARKTECHPOST
Ubie
cogitativo
AQEMIA
somnoware
HEALTH [at] SCALE TECHNOLOGIES
CODOXO
Braid
medicus.ai
Medial EarlySign
nova
AdaptX
hearX
Mednition
mmg
Jubo
KeenEye
aetherAI
HUMA.AI
AIDOT
HISTOWIZ
healx
VERSEON
nucleai
Virgo
Niramai
qure.ai
Lynx.MD
oda
Presagen
Nostos Genomics
neuro42
JuliaOmix
KINTSUGI

ORGANIZATION NAME	WEBSITE	REGION	DESCRIPTION
AKASA	www.akasa.com	USA	AKASA is an AI-powered automation company for revenue cycle management in healthcare.
ConcertAI	www.concertai.com	USA	ConcertAI is an AI-powered SaaS data company in healthcare that develops medical research tool in oncology.
Harrison.ai	www.harrison.ai	USA	Harrison.ai is a clinician-led healthcare AI company using deep learning to improve the odds of IVF success for thousands of patients.
Digital Diagnostics	www.digitaldiagnostics.com	USA	Digital Diagnostics is an AI diagnostics company that transforms the quality, accessibility, and affordability of healthcare through the application of technology in the medical diagnosis and treatment process.
Aizon	www.aizon.ai	USA	Aizon is a cloud-based SaaS platform that helps optimize industrial processes in biotech and pharma with the use of advanced analytics, artificial intelligence, and other smart factory technologies.
Qventus	www.qventus.com	USA	Qventus optimizes operational decisions in hospitals in real time to reduce costs, improve quality and experience using machine learning and optimization algorithms.
ClosedLoop.ai	www.closedloop.ai	USA	ClosedLoop uses the power of AI to identify at-risk patients and recommend the best interventions for each individual.
Enlitic	www.enlitic.com	USA	Enlitic builds medical deep learning products to streamline radiologists' workflows and improve healthcare diagnosis.
Curai Health	www.curaihealth.com	USA	Curai Health is a virtual care company that uses artificial intelligence to provide chat-based primary care at a lower cost.
Redica Systems	www.redica.com	USA	Redica Systems Empowers the champions of regulatory, quality and safety with actionable data intelligence to improve compliance, increase product quality, and build a more efficient organization.
Cleerly	www.cleerlyhealth.com	USA	Cleerly is a healthcare company whose mission is to create digital care pathways to prevent heart attacks by integrating quality clinical science with the latest-generation AI.
Aiforia	www.aiforia.com	Finland	Aiforia's mission is to provide global access to AI-powered image analysis for medical, healthcare professionals and beyond.
Alpha Medical	www.helloalpha.com	USA	Alpha Medical is a digital platform that offers healthcare and telemedicine services online.
Gradient AI	www.gradientai.com	USA	Gradient AI is a provider of Artificial Intelligence and Machine Learning based solutions designed specifically for the insurance industry.

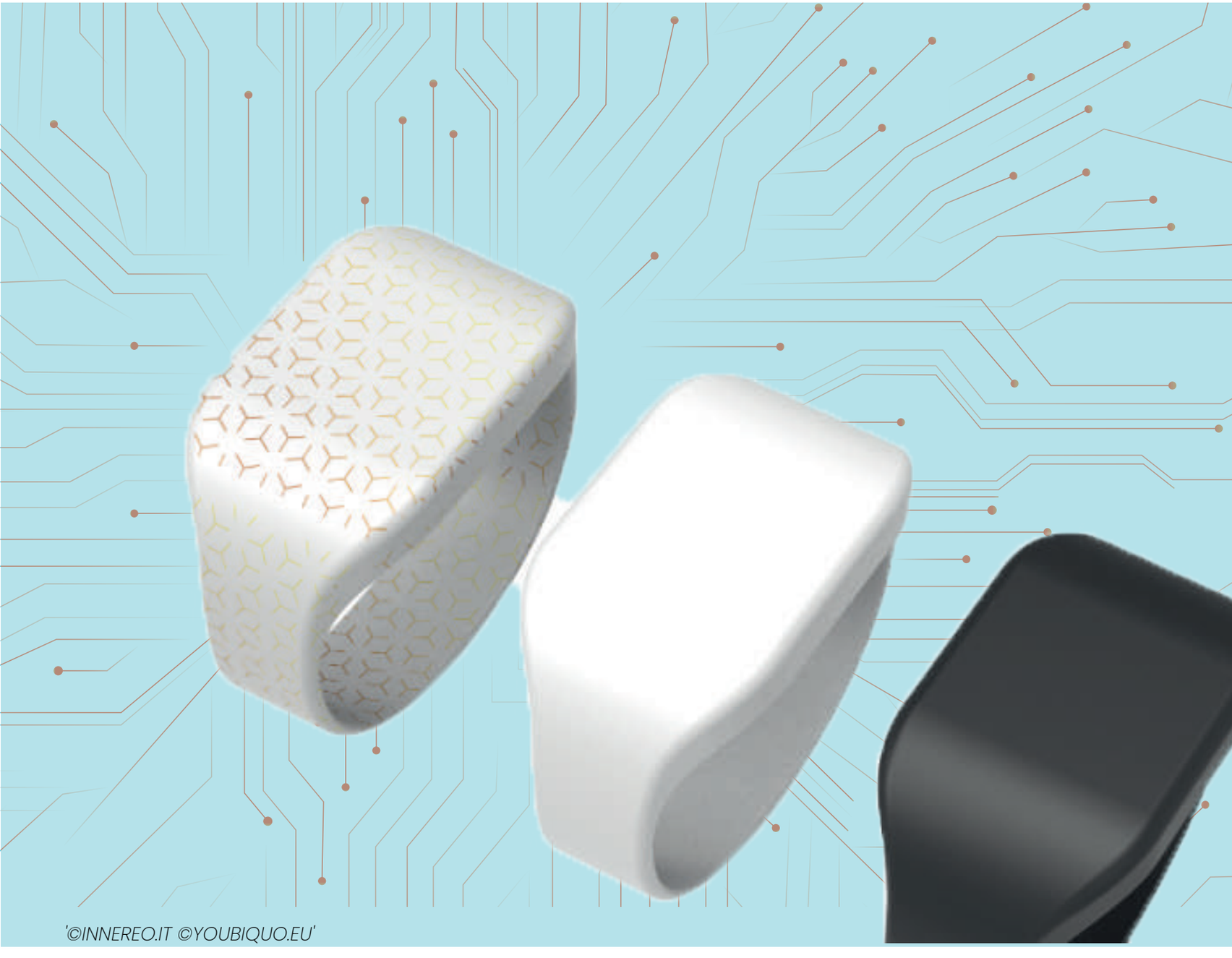
ORGANIZATION NAME	WEBSITE	REGION	DESCRIPTION
Synapsica Healthcare Inc.	www.synapsica.com	India	Synapsica is a B2B health-tech firm that provides AI-enabled automation of diagnostic radiology workflow and reporting.
Healthily by Your.MD	www.livehealthily.com	UK	Healthily is the world's first medically approved self-care app designed around you.
Syllable	www.syllable.ai	USA	Syllable uses AI to transform the way healthcare systems initiate and cultivate relationships with patients.
Diagnostic Robotics	www.diagnosticrobotics.com	Israel	Diagnostic Robotics develops a signal-agnostic artificial intelligence system for healthcare insurers, providers, and patients.
Unlearn.AI	www.unlearn.ai	USA	Unlearn.AI uses AI to create Digital Twins and is a platform to accelerate clinical trials for Alzheimer's Disease, Multiple Sclerosis, and other complex diseases.
Reverie Labs	www.reverielabs.com	USA	Reverie Labs is a machine learning-driven pharma company for accelerating preclinical drug development.
Infermedica	www.infermedica.com	Poland	Infermedica is a leading digital health company specialized in AI-powered solutions for preliminary diagnosis and patient triage.
Pieces	www.piecestech.com	USA	Pieces is a healthcare artificial intelligence and technology company that connects health systems and the community to address clinical and social determinants of health through community networks and intelligent software and services.
Orbita	www.orbita.ai	USA	Orbita provides software to improve patient engagement in digital healthcare through voice-first and conversational AI solutions.
Prealize Health	www.prealizehealth.com	USA	Prealize Health is a predictive analytics company helping health plans, providers, and employers transform healthcare from reactive to proactive, so that more people live healthier lives.
Lumiata	www.lumiata.com	USA	Lumiata driving predictability and precision at scale—empowering healthcare organizations to reduce risk and manage costs with AI.
Ubie	www.ubie.life/en	Japan	Ubie, a Tokyo-based health-tech startup, provides AI-based healthcare products, including a hospital SaaS product and an AI symptom checker.
Vara	www.vara.ai	Germany	Vara develops a breast cancer screening platform powered by AI for radiologists.
Cogitativo	www.cogitativo.com	USA	Cogitativo is a machine learning enterprise offering world class solutions to hospitals, physicians, and payers.

ORGANIZATION NAME	WEBSITE	REGION	DESCRIPTION
HEALTH[at] SCALE Technologies	www.healthatscale.com	USA	Health at Scale is a healthcare machine intelligence company that uses proprietary advances in artificial intelligence and machine learning to optimize care delivery for individuals by empowering at-risk payers, employers and providers.
Somnoware Healthcare Systems	www.somnoware.com	USA	Somnoware is transforming respiratory healthcare across the globe by reducing cost and improving care management.
Codoxo	www.codoxo.com	USA	Codoxo uses AI technology to re-envision healthcare affordability and effectiveness, including a patented algorithm to identify problems and suspicious behavior earlier than traditional techniques which helps ensure our scarce healthcare dollars go to real patient care.
Braid Health	www.braid.health/www	USA	Braid Health develops an AI-powered diagnostic collaboration platform that supports radiologists and healthcare providers.
Medicus AI	www.medicus.ai	Austria	Medicus AI designs and develops an artificial intelligence (AI) based enterprise and app platform that bring meaning to health data.
Medial EarlySign	www.earlysign.com	Israel	Medial EarlySign helps healthcare clients with their early detection and prevention of high-burden diseases using routine, existing EHR data
Novadiscovery	www.novadiscovery.com	France	Health tech company using clinical trial simulation technology to predict drug efficacy and optimize clinical trial development.
AdaptX	www.adaptx.com	USA	AdaptX is a clinical management platform that delivers analytics and AI to hospitals and surgery centers.
hearX Group	www.hearxgroup.com	South Africa	Affordable access to hearing care using digital solutions that anyone can use, anywhere
Mednition	www.mednition.com	USA	Mednition offers a machine learning-powered decision support solution for hospitals.
MMG	www.mmg-ai.com/en/	Spain	MMG develops A.I. solutions that allow individuals and institutions to access, discover, and exchange knowledge to make better decisions.
Jubo	www.jubo-health.com	Taiwan	We deliver person-centered care experience by connecting healthcare data, empowering caregivers, and engaging family.
Keen Eye Technologies	www.keeneye.ai	France	Keen Eye Technologies is an AI platform company that elevates computational pathology to guide better drug development and diagnostics.
aetherAI	www.aetherai.com	Taiwan	aetherAI is a medical image AI company that focuses on improving digital pathology.
AIDOT	www.aidot.ai	South Korea	AIDOT develops medical software using AI. Their flagship product makes cervical cancer screening faster, easier and more accurate.

ORGANIZATION NAME	WEBSITE	REGION	DESCRIPTION
Huma.ai	www.huma.ai	USA	Huma.ai is a business intelligence platform built to automate healthcare through human language: What if you could just ask questions of your data and instantly get answers - that is Huma.AI.
HistoWiz	www.home.histowiz.com	USA	HistoWiz provides automated histopathology for biomedical researchers.
Healx	www.healx.io	UK	Healx is an AI-powered, patient-inspired technology company, pioneering the next generation of drug discovery for rare diseases.
Verseon	www.verseon.com	USA	Verseon's breakthroughs in computational chemistry, AI, & synthesis create novel pharmaceuticals. Atom by atom, we drive innovation.
Nucleai	www.nucleaimd.com	Israel	Nucleai is a spatial biology company that has developed an AI-powered platform to enhance drug development and improve patient outcomes.
Virgo Surgical Video Solutions	www.virgosvs.com	USA	Virgo accelerates clinical trials with AI for endoscopy
NIRAMAI Health Analytix	www.niramai.com/technology/	India	NIRAMAI has developed a novel breast cancer screening solution that uses Thermalytix.
Qure AI	www.qure.ai	India	Qure.ai builds deep learning solutions that aid physicians with routine diagnosis and treatment, allowing them to spend more time with patients, and to make healthcare more affordable and accessible using the power of AI
Ada Health	www.ada.com	Germany	Ada Health is a digital health company that helps to diagnose symptoms and offer treatment advice using an AI-enabled platform.
Lynx MD	www.lynx.md	USA	Lynx MD is a world where the clinical data can be shared securely.
Presagen	www.presagen.com	USA	Presagen is The Social Network for Healthcare, connecting clinics, patients and medical data from around the world through AI.
Nostos Genomics	www.nostos-genomics.com	Germany	Nostos Genocmics' easy-to-use, AI-driven platform AION is changing the way labs can interpret genetic test results. So patients can get fast and accurate insight.
Neuro42	www.neuro42.ai/#page1	USA	Neuro42 is developing and commercializing advanced MRI technology for diagnosis and image-guided interventions.
GenomeUp	www.juliaomix.com/about-us/	Italy	GenomeUp is an AI platform for enabling doctors and clinical institutions to diagnose and treat rare genetic diseases in less than 24 hours.
Kintsugi	www.kintsugihello.com	USA	Kintsugi is an AI-based platform that detects clinical depression and anxiety using machine learning and voice biomarkers.
Aqemia	www.aqemia.com	France	Agemia is a next-gen pharma tech company using unique quantum and statistical mechanics algorithms to fuel a generative artificial intelligence to design novel drug candidates.

An Italian AI Startup Turning Emotional Unconsciousness To Instant Music

By: Innereo Team (Italy)



'©INNEREO.IT ©YOUBIQUO.EU'

Sebastiano- CEO

The head of the company is Sebastiano Deva, CEO, digital entrepreneur, and philosopher.

Emotions play a fundamental role in human life, influencing the mental and physiological processes of our species. Emotions can be defined as complex response configurations, selected during the course of evolution to favor the adaptation of the organism to the environment, from which stimuli or representations are received that upset its equilibrium. As response mechanisms, emotions often involve similar neurophysiological and biochemical modifications, assuming a social and relational significance within the species. In other cases, however, they may manifest themselves differently, modulating according to the subjective experiences of each individual.

Affective computing, sometimes also referred to as Artificial Emotional Intelligence, is the branch of Artificial Intelligence (AI) that develops technologies able to recognize and express emotions. In virtue of this new perspective where classic artificial intelligence is integrated with emotional intelligence, we now speak of emotional AI or the combination of emotional and artificial intelligence. Advances in affective computing technology have led to the growth of emotion recognition research in recent years. Systems able to perceive emotions bring multiple benefits to their users, as they are useful both to the user, who becomes more aware of the emotions he or she is showing, and to developers, who can make use of emotion recognition to make their projects adaptive to the user's experience, as well as to support the detection of cognitive disorders, anxiety, or stress. This can be found both in the health field and in the work environment, where negative emotional situations can affect the productivity of employees.

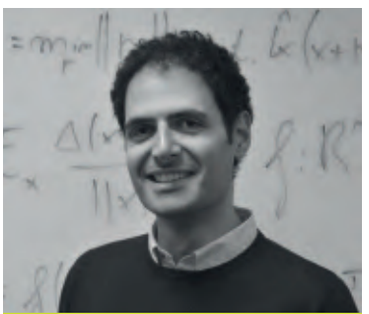
At Innereo, we believe that the first step towards a reduction of stress levels goes through the understanding of the feelings we are experiencing in a given moment. To that aim, we designed an affective computing system able to analyze a subject's emotional state, transliterating it to an emotional experience that allows people to feel their feelings through their senses. One possible implementation of this concept is DeepSoundMe, a mobile application that analyzes the Galvanic Skin Response (GSR) to create music in real-time, based on your feelings. To this aim, together with YOUBIQUO we realized a brand-new wearable device able to catch different biometric signals (including GSR) while ensuring long battery life and a very high wearing comfort.

The GSR is widely used to recognize emotions. However, the signal itself is not trivial to analyze. The core of our system is an AI module designed and trained to cope with this high variability, making it possible to

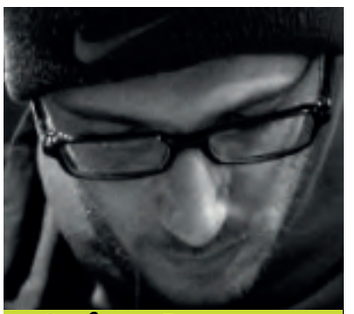
obtain real-time subjects' emotional shades from the GSR values. The AI module consists of an innovative deep neural network able to disentangle the emotional-related information from those associated with external and not of interest, factors. As the music is generated based on feelings, each time the felt experience is new and different, allowing people to reconnect with their inner ego. In fact, thanks to the instant evaluation of the emotional state by the AI process, the user can immerse himself in a sound experience of "Healing Music", generated by his own emotional patterns. Starting from the instantaneous detection of GSR, the sound experience results from the combination of more than 33 thousand musical patterns.

Innereo Team (Italy):

DeepSoundMe is part of Innereo, an Italian start-up recently founded by five peers, each expert in a given domain. The head of the company is Sebastiano Deva, CEO, digital entrepreneur, and philosopher. The research and innovation strategy is led by Stefano Marrone, CTO, Ph.D. in Information Technologies and Electrical Engineering, research fellow, and lecturer at the University Naples Federico II. He authored more than 30 papers published in international journals and proceedings on AI and pattern recognition, with applications ranging from biomedical to biometrics, security ethics, and fairness. Public relations and Marketing are handled by Massimo Morgante, CMO, a digital entrepreneur with a long start-up founding experience, while the software design and development is led by Stefano Pallozzi, SD, computer systems designer and analyst for several tech companies. All the music generation, study, and design are made by Pier Paolo Polcari, a music designer and composer known at a national and international level. Recently, we had a guest collaboration with Gloria Cretella, computer science engineer, Apple Developer Academy graduate, and three-times Apple WWDC winner.

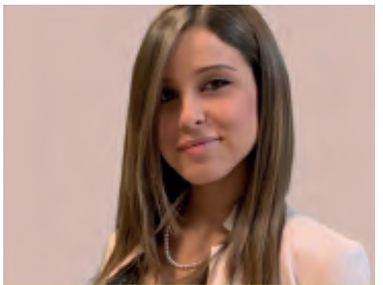


Stefano M.- CTO



Stefano P. - Senior Developer

Massimo- CMO



Gloria AI Developer

Paolo - Sound Composer

"Wearable Hardware project is led by YOUBIQUO (Italy)"

ANALYSIS OF REDDIT SELF-REPORTED DEPRESSION DIAGNOSIS (RSDD) DATASET

Reddit consists of large open-source data on self-reported depression. RSDD is an uncurated text-based dataset, to enable the detection of depression. A common approach to creating a curated dataset is by web-scraping with fixed labeling rules (AlSagri and Ykhlef, 2020; Cornn, 2019; Shen et al., 2017b). The problem with such a dataset is the presence of labeled noise, that occurs because of mislabeling by non-experts or oversimplified labeling criteria. Oversimplified labeling criteria may lead to mislabeling of a non-depressive statement in the /depression subreddit as depressed (Cornn, 2019). Due to the pattern-memorization effects, label noise may significantly compromise the performance of deep learning models in classification tasks (Flatow and Penner, 2017; Zhu and Wu, 2004), particularly in the detection of depression (Cornn, 2019).

Large datasets suffer from noisy labels, due to erroneous automatic and human annotation procedures. Language-based tasks, recent studies addressed the issue of label noise by supplying additional contextual information to the attention models. The attention mechanism individually computes the attention weights of each token over the bag-of-words tokens (Vaswani et al., 2017). As a result, attention models such as generative pre- trained transformers (GPT) (Radford et al., 2019) and BERT (Devlin et al., 2018) neglect the contextual information in the calculation of dependencies between tokens (Yang et al., 2019a).

In the present study, the deepkapha AI Research Lab's **Project Depressio** team scraped web-scrape posts from January 2018 to November 2020 in the 2 subreddits /depression and /AskReddit, which correspond to "depressed" and "non-depressed" classes, respectively (see Table 1). They exploited the contexts for robustness to label noise in the detection of depression where two data augmentation methods, i.e., Negative Embedding and ***Empath (Chen and Bernstein, 2016)*** were proposed. This method was not proposed earlier in other studies related to depression.

Fast, E., Chen, B., & Bernstein, M. S. (2016, May). Empath: Understanding topic signals in large-scale text. In Proceedings of the 2016 CHI conference on human factors in computing systems (pp. 4647-4657).

/Depression		/AskReddit	
Title:	Comments:	Title:	Comments:
I am so tired of people taking me for granted.I give them too much of energy. I am sick of everything. my life, my family, my friends	- I'm sorry. I'm really hoping the best for you. - I know how you feel. I feel exactly the same right now. I wish I could give this post a thousand rewards.	What's something that impresses most people that doesn't impress you?	Limousines. As a kid, I used to think that was the sign that you made it. Now I realize you just need $95 - If you've get more than 5 people getting a limo or party bus is miles cheaper than getting multiple Ubers. Plus you can drink in them.

TABLE 1: Samples of the Reddit self-reported depression diagnosis (rsdd) dataset for /depression and /askreddit. the first comment in /depression is a non-depressive sentence. This is an example of label noise.

Original Text	Lexicons	Post-processed Texts
Wow. I understand that the rules are the rules, you just painted "everyone" who offers that as either a psycho or a predator. I must say I am feeling like one now because ...	hate, nervous, suffering, art, optimism, fear, zerst, speaking, sympathy, sadness, joy, lust, shame, pain,negative emotion, contentment, positive emotion, depression, pronoun, ...	Wow. I understand that the rules are the rules, you just pained "everyone" who offers that as ... hate, nervous, suffering, art, optimism, fear, zest, speaking, sympathy, sadness, joy, lust, shame, pain,

TABLE 2: Example of empathy generating lexicons and concatenating generated lexicons with the original text.

What is Empath?

The context of a sentence can be exploited by generating high-level lexicons which represent the overall emotional context. Researchers have relied on such high-level lexicons to identify signs of depression in social media posts and to understand the overall meaning of texts at scale. One of the most commonly used libraries is Linguistic Inquiry and Word Count (LIWC) which counts words relevant to lexical categories such as sadness, health, and positive emotions (Tausczik and Pennebaker, 2010). For example, positive lexicons include words such as happiness, joy, fun etc. LIWC has a fixed list of 40 lexical categories that limit its ability to capture signs of depression in text data. Empath library is designed using deep learning techniques and crowdsourcing that allow it to in- corporate new lexical categories (Fast et al., 2016). In the present study, the proposed data augmentation method Empathy utilizes the Empath library and initially updates the library with 2 lexicons, "pronoun" and "depression", which consider relevant words as possible indicators of depression. This process is theoretically aligned with previous findings that depressed patients use first-person singular pronouns and depression-related words more frequently than healthy controls (Rude, Gortner, and Pennebaker, 2004). Each text sample is evaluated by the Empath library to generate high-level lexicons which are then linearly concatenated with the text sample into a new text sample (see Table 2). The generated text sample consists of both original contexts and high-level, extracted emotional contexts.

What is Negative Embedding?

Original text:	This is so frustrating. I'm sorry you're experiencing this. I know how you feel.																									
Tokens:	[CLS]	this	is	so	frustrat	##ing	.	[SEP]	i	'm	sorry	you	're	experienc	##ing	this	.	[SEP]	i	know	how	you	feel	.	[SEP]	
Segment Embeddings:	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	
Negative Embeddings:	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Figure 1: A comparison between Segment Embeddings and Negative Embeddings. In Segment Embeddings, numbers represent sentence indices. In Negative Embeddings, 1 represents negative words and 0 represents non-negative words

In this work negative embeddings are used in order to emphasize depressive contexts on conditions of the existence of negative tokens and to estimate true labels from noisy labels. The negative embedding labels binary classes (1 and 0) for negative and non-negative tokens respectively (Figure 1). The objective of negative embedding is to compute the probability of depression p(·|S\Σ ni) given the negative token ni, the sequence S and the context S\Σni. The negative tokens are common negative tokens in the sentiment analysis task and have been pre-defined in previous studies (Hu and Liu, 2004; Liu et al., 2005).

BERT and DistilBERT

Bidirectional encoder representations from Transformers (BERT) (Devlin et al., 2018) is a widely used language model that researchers extensively implemented to achieve state-of-the-art performance in various language understanding tasks. As BERT is composed of attention-based transformer blocks and pre-trained on large corpora (Zhu et al., 2015), it can capture a variety of linguistic features and contexts.

BERT and DistilBERT were tested with Empath and Negative Embeddings the two data augmentation techniques. Empathy, we firstly added 2 depression-related lexicons (“pronoun” and “depression”) to the Empath library and applied the library to analyze text and generate emotional lexicons. The lexicons were then concatenated with original texts into new text samples (Figure 1), which are finally tokenized by the WordPiece tokenization (Wu et al., 2016). The codes used in this study have been made publicly available2.

The performances of BERT and DistilBERT with Negative Embedding is around ~85%. BERT and DistilBERT with Negative Embedding converge faster than BERT and DistilBERT without text augmentation.

GITHUB

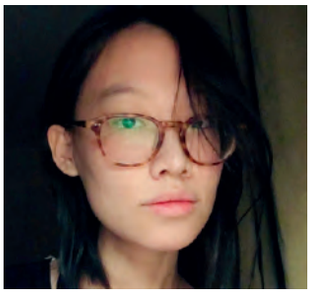
https://github.com/deepkapha/depressio/tree/restructure

Author Details

Tannistha Maiti

Dr. Tannistha Maiti is a PhD in Geophysics and Seismology from University of Calgary. She has worked on various topics that ranges from Earth Science to computation and mathematical modelling. She is passionate about machine learning applications in the Energy and healthcare sector. She has extensively worked in mathematical modelling and computational Geophysics during her PhD at U Calgary and MS studies at Virginia Tech. She also holds an undergraduate degree from the prestigious university at IIT Kharagpur, India. In her research career spanning for about 10 years she has been involved in various projects and has more than 20 peer reviewed publications in various conferences and journals. She also has extensive teaching experience and is passionate to share the knowledge with peers and newcomers. During her PhD studies she received various accolades for her work. She leads MLOPs engineering practice within DK AI Labs, participates in various research projects and supervises several deep learning internships within deepkapha AL Lab. In her spare time, she enjoys blogging both technical and non-technical.

Dr. Jie Mei



Dr. Jie Mei received her M.S. in cognitive neuroscience from École Normale Supérieure Paris in 2015, and completed her Ph.D. in medical neuroscience at Charité – Universitätsmediz in Berlin in 2019. Her major research interests include neuro-inspired AI, computational neuroscience, and application of machine learning in neuroscience and neurology. She is Head of AI Research at DeepKapha AI Research.

Aninda Bhattacharjee

Aninda Bhattacharjee is a Big Data Engineer by profession who has about three years of experience in the field of innovative technologies. An individual highly passionate about emerging technologies which make deploying and scaling ML applications easy and thus production ready. His interests lie in the NLP domain and he loves to experiment with various graph-based algorithms and architectures in his spare time.

Dat Ngo

Dat Ngo is software intern at Meta AI. He is currently completing his Master in Computer Science - Intelligent System with focus on AI, NLP, and Graph Representation Learning from UT Dallas. For industry experience, he has interned as SWE, ML engineer, and AI researcher at deepkapha AI Research, Samsung, 7-Eleven, and Meta (formerly Facebook). He has worked as Visiting AI Researcher at deepkapha AI Lab.

Spotlight Interview with

NATALIA VASSILIEVA

Understanding The Cerebras High-End Compute Power And Role in AI and Healthcare: Talk with Natalia Vassilieva Director of Product, Machine Learning at Cerebras Systems

Natalia Vassilieva is Director of Product, Machine Learning at Cerebras Systems, a computer systems company dedicated to accelerating deep learning. Her focus is machine learning and artificial intelligence, analytics, and application-driven software-hardware optimization and co-design. Prior to joining Cerebras, Natalia was a Sr. Research Manager at Hewlett Packard Labs, where she led the Software and AI group and served as the head of HP Labs Russia from 2011 until 2015. Prior to HPE, she was an Associate Professor at St. Petersburg State University in Russia and worked as a software engineer for several IT companies. Natalia holds a Ph.D. in computer science from St. Petersburg State University.

TELL US ABOUT CEREBRAS SYSTEMS.

Cerebras is an AI systems company. We've built a new type of computer system to greatly accelerate the training of deep neural networks (DNN) and open up new areas of research so that scientists and practitioners can do previously impossible work.

CAN YOU TELL US A LITTLE ABOUT YOUR ROLE AS DIRECTOR OF PRODUCT, MACHINE LEARNING, AT CEREBRAS?

My role is rigorous. As a Director of Product, I sit between our engineers, customers, and the market overall. My role is to understand what kind of product we should be building, how it's useful for our customers, and how it can potentially open doors in other markets. In practice, what that means is looking into trends in the industry overall. In terms of AI, we need to keep an eye on the latest research to understand where the field is going. Machine learning is a very fast-evolving field and many new research papers are published daily.

Once new research is published, typically with some delay, enterprises adopt these new methods to make it easier for them to use in hardware and other applications. We are looking at state-of-the-art research, what customers want to do with that research, and what kind of applications our customers are seeking to solve. We ask, what kind of methods can be applied to help them with their task? Collecting data on the engineering organization's requirements enables us to develop the next version of our product or software release.

CAN YOU TELL US A BIT ABOUT YOUR PRODUCTS, SPECIFICALLY THOSE WITH APPLICATIONS IN HEALTHCARE, PHARMA, AND DRUG DISCOVERY?

At Cerebras we built the world's largest and fastest AI computer – the ***CS-2 system***. It is a very powerful computer that enables you to train deep neural networks in hours or days vs the weeks or months it takes with legacy hardware. What we're hearing from our customers is that when you're working on cutting-edge research, time matters. Being able to train a model in hours or days means that researchers can test many more hypotheses that can lead to major scientific breakthroughs.

For example, we are working with pharmaceutical leader GlaxoSmithKline to use AI for drug discovery. They had a hypothesis that adding epigenomic data to their AI models would lead to more accurate and useful models. But they were previously unable to test this hypothesis because it would take too long to run on legacy hardware. They called us and we got them onto our CS-1 system. They were able to prove their hypothesis that by adding epigenomic data they could improve their models.

In regards to the pharma industry, there has been a rise in the quality of models when applied to modeling sequence data. You can think about the natural language of text as a sequence of characters or a sequence of words. People find out how to train efficient and representative models on that data in a self-supervised manner, where you don't need any labels. You just feed all the text that you have, and it learns representations and can do some useful tasks for them. Many models have been designed to represent natural language and sequence data. The models created for language are directly applicable to modeling for biological tasks.

There is growing interest in working in domain-specific text. Being able to get insights from medical literature and to understand what kind of information can be derived from clinical reports or from any written text is important. In biotech, there are many examples of sequence data. Some examples of biological sequences include proteins, the sequence of amino acids, and DNA. If you want to model what happens in the genome, it's a lot of modeling those sequences.

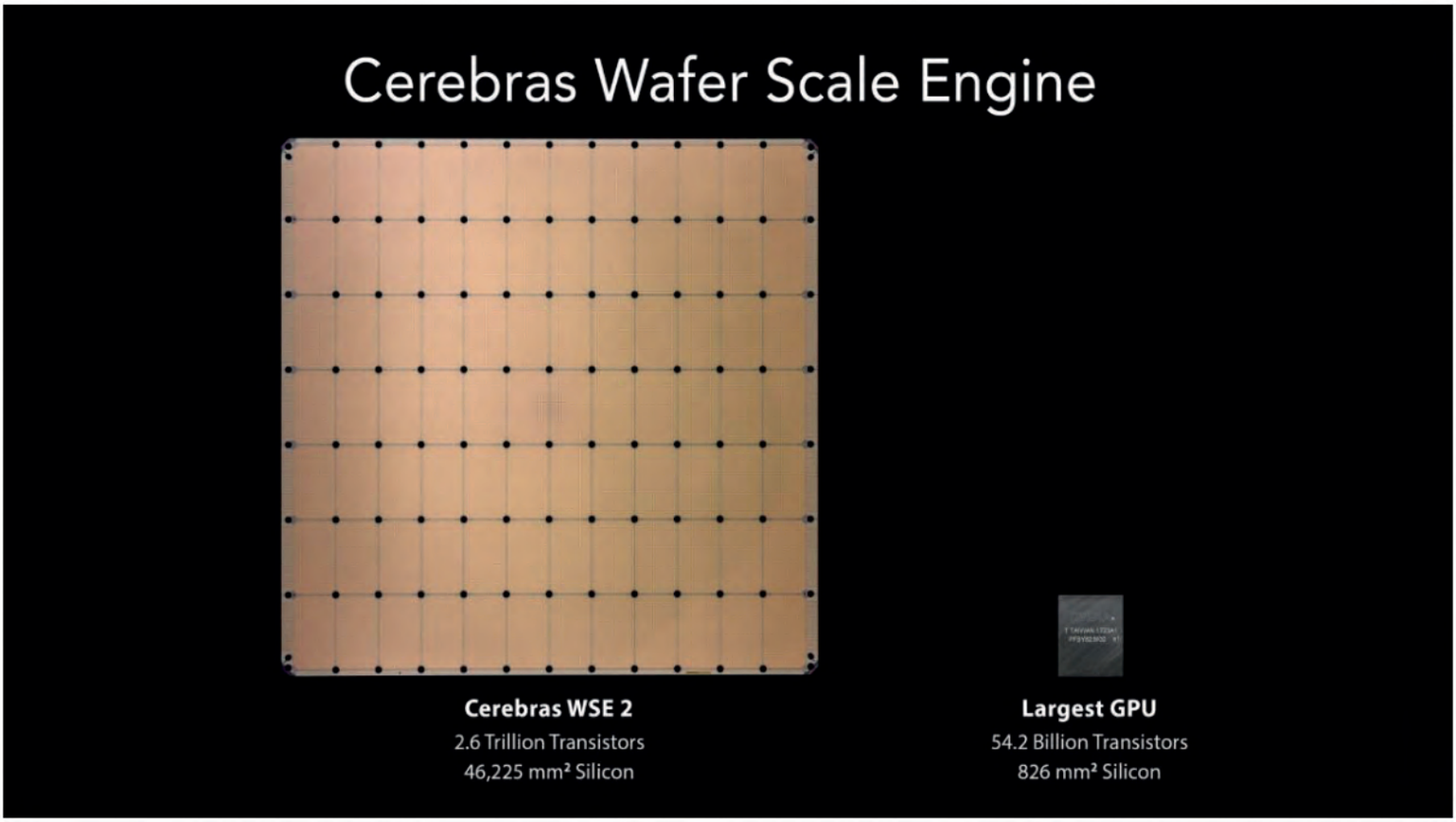
These models are typically quite compute-intensive. High compute-intensive tasks require a heavy infrastructure footprint to be trained in a reasonable time. It is challenging to train at high scale on existing, traditional hardware. Our hardware is capable of accelerating the training of those types of models significantly. We are relevant to pharma because of our ability to process data faster with the CS-2 system.

WHAT IS THE CEREBRAS CS-2 SYSTEM? HOW DOES CEREBRAS USE AI TO DRIVE FASTER DRUG DISCOVERY? HOW DOES THE CS-2 DIFFER FROM YOUR COMPETITORS?

The Cerebras CS-2 is our second generation system. While our competitors are trying to connect together many weak processors, we have one giant processor that can train very large models. One of the main innovations in the CS-2 is the Wafer-Scale Engine which has 850,000 cores. This is significantly more than you can find on any CPU or GPU. It gives us the ability to significantly accelerate tasks that will require a lot of computing power.

With traditional hardware for high com-

pute-intensive tasks, researchers are forced to cluster or connect together multiple traditional processors to be able to complete work in a reasonable time. It's not very efficient. Instead of connecting multiple, small core count processors, we can use our single, large chip. It is easier to leverage the computing power of many cores, all packaged in a single device. The CS-2 system accelerates different compute tasks, such as the training of deep neural networks.



WHAT ARE SOME OF THE BIGGEST CHALLENGES THAT CEREBRAS IS LOOKING TO ADDRESS IN HEALTHCARE AND OTHER VERTICALS?

Across all verticals, the main value proposition we offer is a powerful tool that allows domain specialists to complete their experiments much faster. We want to enable researchers to learn more quickly from the results of their experiments.

The field of machine learning is, by and large, a field of trial and error. There is no golden book of rules on how to compose a specific model. Typically, you need to try many different things before converging on something that works for your problem. The speed of experimentation and the speed at which you can run those different trials is extremely important.

With our hardware, we give researchers a way to make those trials take less time. We let them test many more hypotheses than they would be able to do otherwise. What we often find in practice is that researchers start with the one or two ideas they want to test. It often takes months to test a single idea in traditional environments. With Cerebras, you can test more ideas, and test them faster.

The reality is if you don't have the result of those experiments, it kind of slows down your imagination. If a tool can get some results in a matter of hours or days, the number of ideas that the researchers generate just explodes. Once a researcher can see what works and what doesn't work, they can come up with 2-5 new ideas they want to test out. It fuels creativity and accelerates research significantly.

CAN YOU TELL US ABOUT CEREBRAS'S LATEST PARTNERSHIPS IN HEALTHCARE AND AI?

We have several projects. Our partnerships with pharma companies provide a tool that enables them to create and develop new AI-driven methods. In the case of GlaxoSmithKline (GSK), we are helping them on the path to new therapeutics and new vaccines, while getting insights with the help from artificial intelligence along the way.

Another example is the collaboration with ***AstraZeneca***. AstraZeneca has been interested in developing an internal search engine that will enable a question and answering engine. This Q&A engine will allow their researchers to find where to quickly access answers to questions about past research and past clinical trials. Another task has been building a domain-specific language model, which can help them build the question answering and machine translation engines.

HOW DOES THE CEREBRAS PLATFORM GIVE VALUE TO ITS CUSTOMERS?

Typically in healthcare, we work with computational chemists, experts in biology and bioinformatics. Many of them are experts in machine learning, but almost none are experts in distributed programming. It really should be easy for them to test their ideas without knowing how the hardware works underneath, and without spending too much time thinking about they should optimize certain tasks. There is great value in running experiments

much faster and making it easier for the researchers. Ease of use and fast experimentation is critical. And that is what our system brings to the table.

I am from Russia, so let me share one more analogy from my university days. My first programming classes were taken when we were allowed just one hour on a computer. You needed to complete all your programs on a piece of paper first. You got one chance to test if your program runs right. You had to think really carefully about how you design that program, how you write that down, and then you either get it right or not and you don't have any other chances. In many cases right now, researchers are in the same situation with these deep neural networks. When it takes you months to test your hypotheses, you know that you have only one shot, and it limits what you can do.

Our system has essentially reduced the cost of curiosity. It enables you to not have to spend so many resources on checking whether your idea is worth pursuing. I can go ahead and test it and get more insight faster.

Using AI to study MRI Scans and Understand Depression

Major depressive disorder (MDD), commonly called depression is ranked by WHO as the single largest contributor to global disability and majorly contributes to suicide deaths. "***At a global level, over 300 million people are estimated to suffer from depression, equivalent to 4.4% of the world's population.***" The illness is still under-recognized and under-treated despite the alarming rate at which it is affecting the population and is predicted to be the leading health burden worldwide by 2030.

A person suffering from depression is likely to experience sadness, low self-esteem or guilt, diminished interest in normal activity, fatigue, and disturbances of sleep or appetite. It can either manifest as chronic illness or appear as short episodes that hinder an individual's ability to function normally at work or cope with daily life. Clinical criteria used to diagnose depression have traditionally included a systematic review of the symptoms and history of the patient, usually gathered through questionnaires and interviews. These behavioral observations form the basis of the diagnosis and can therefore lead to methodological errors. Additionally, depressive symptoms often mimic other disorders and can create confusion in the diagnosis. It is imperative that we standardize the methods and techniques used to gather and analyze the data in order to resolve this issue.

In order to develop clinically useful information about MDD for diagnosis and prognosis, there is a pressing need to map our current understanding of the brain dysfunction involved. The field of neuroimaging has undergone tremendous advancements in recent years, allowing for the identification of measurable indicators related to several mental disorders. Although progress has been made in the area of finding neuroimaging markers for MDD, the progress has stagnated due to limited data, underpowered studies, and failing to replicate findings. It is possible to employ findings from brain imaging studies of the regional structure and functional connectivity of the brain for solving this problem. This requires a method with a rigorous analytic approach that can deal with large feature spaces and relatively few observations.

In tandem with the development of computational science, machine learning algorithms are now proven to be powerful tools for processing high-dimensional datasets and accurately identifying disease features which aid in clinical diagnosis. ***A recent study demonstrated a deep neural network model that predicted skin cancers with a similar precision to that of dermatologists***. The individualized prediction and characterization of patients with psychiatric disorders has made groundbreaking advances in machine learning techniques.

Methodology

Deepkapha AI Research Lab's Project Depression team analyzed structural MRI scans using machine learning algorithms to categorize depressive and control patients. We acquired the data for patients with depression (51 participants) and healthy participants (21 participants) on the open-source platform, ***OpenNeuro***. This data was collected using the standardised Brain Imaging Data Structure (BIDS) format, which saved time rearranging files.

In order to derive individual-level measures from the dataset, we used harmonised imaging processing protocols such as ***FreeSurfer*** and then used them for regression analysis. This open-source software performs a variety of functions, such as processing MRI data and measuring morphometric properties that are ultimately used to develop a computerized model of the brain. The morphometric data for multiple volumetric T1 weighted images was generated by cortical recon-

struction. We used version 7.1.1 of FreeSurfer in this study.

We processed the data using the standard 'recon-all' pipeline, which included correction for motion, removal of non-brain tissue, automated Talairach transformation, intensity normalization for the surface, and segmentation of the cortex and subcortex by intensity, along with statistical analysis to determine differences between groups. The ***Desikan-Killiany Atlas*** was used to parcellate cortical surfaces. These parcellations of subcortical regions were then used to generate statistical data corresponding to the surface area, volume, and thickness of the subcortical regions. Using this approach, we were not only able to overcome the small sample size but also better analyze the study-specific covariates.

A total of 34 brain regions were identified for each participant across all three modes of measurement. The volumetric measurements in cross-sectional studies are not used in their original form, but are corrected for the volume of the cranium, also called intracranial volume (ICV). A volume correction is necessary in order to ensure that the changes that are being experienced are not caused by other factors such as gender differences or different image acquisition protocols. Following that, each of the features was paired with its risk of depression, and the results were then sorted. For the first part of the study, we used statistical tests such as t-tests in order to show whether groups differed on a set of extracted features.

In addition to using univariate group-level statistical tests, we used filtering methods such as Random Forest and XGBoost, which allocated scores to each feature and then selected a few of the most prominent ones. For Hyper Parameter Optimization (HPO) we used GridSearch coupled with cross-validation, which produced the best results for the Random Forest classifier.

The feature selection was performed to reduce overfitting on the small dataset. The dataset was split in two train sets with thirty and eleven (depressed and healthy) samples and twenty and ten (depressed and healthy) samples. The preprocessed data is trained and tested on five different classification models, such as logistic regression, SVM, random forest, multilayer perceptron and KNN. For each model, we report the average accuracy and F1 scores across four cross-validations, as well as the best configuration of hyperparameters.

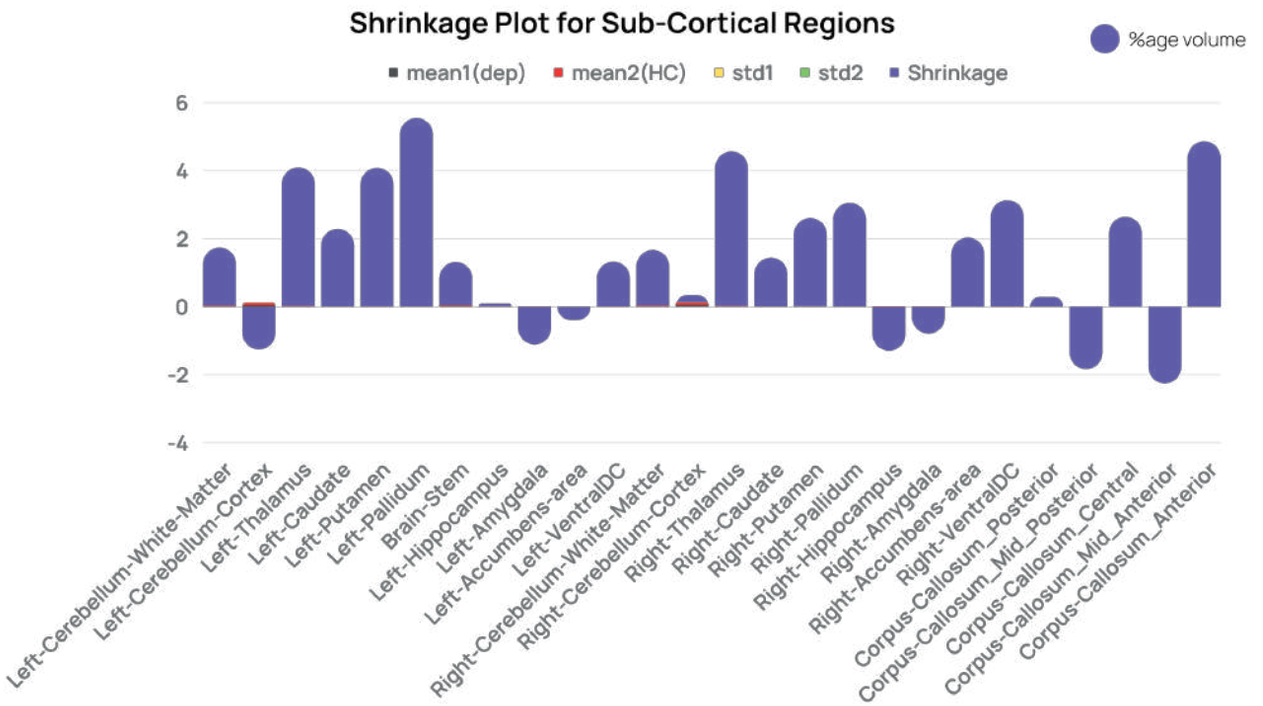


Table 1: Volume

MODEL	LEFT HEMISPHERE	RIGHT HEMISPHERE
K- Nearest Neighbors	We found the accuracy,F1 scores for Train: 0.98, 0.98 Test: 0.70 and 0.70 respectively	We found the accuracy,F1 scores for Train: 0.90, 0.91 Test: 0.50 and 0.49 respectively
Support Vector Machine	We found the accuracy,F1 scores for Train: 0.94, 0.94 Test: 0.63 and 0.61 respectively	We found the accuracy,F1 scores for Train: 0.92, 0.93 Test: 0.50 and 0.51 respectively
Multi-layer Perceptron (best)	We found the accuracy,F1 scores for Train: 0.82, 0.85 Test: 0.73 and 0.73 respectively	We found the accuracy,F1 scores for Train: 0.81, 0.84 Test: 0.73 and 0.72 respectively
Random Forest	We found the accuracy,F1 scores for Train: 0.98, 0.98 Test: 0.73 and 0.72 respectively	We found the accuracy,F1 scores for Train: 0.94, 0.95 Test: 0.63 and 0.52 respectively
Logistic Regression	We found the accuracy,F1 scores for Train: 0.73, 0.76 Test: 0.70 and 0.68 respectively	We found the accuracy,F1 scores for Train: 0.71, 0.74 Test: 0.73 and 0.72 respectively

Table 2: Area

MODEL	LEFT HEMISPHERE	RIGHT HEMISPHERE
K- Nearest Neighbors	We found the accuracy,F1 scores for Train: 0.94, 0.94 Test: 0.60 and 0.60 respectively	We found the accuracy,F1 scores for Train: 0.94, 0.95 Test: 0.67 and 0.66 respectively
Support Vector Machine	We found the accuracy,F1 scores for Train: 0.88, 0.88 Test: 0.67 and 0.64 respectively	We found the accuracy,F1 scores for Train: 0.96, 0.96 Test: 0.70 and 0.68 respectively
Multi-layer Perceptron (best)	We found the accuracy,F1 scores for Train: 0.77, 0.81 Test: 0.60 and 0.60 respectively	We found the accuracy,F1 scores for Train: 0.82, 0.85 Test: 0.77 and 0.75 respectively
Random Forest	We found the accuracy,F1 scores for Train: 0.96, 0.96 Test: 0.60 and 0.57 respectively	We found the accuracy,F1 scores for Train: 0.98, 0.98 Test: 0.60 and 0.59 respectively
Logistic Regression	We found the accuracy,F1 scores for Train: 0.70, 0.72 Test: 0.57 and 0.57 respectively	We found the accuracy,F1 scores for Train: 0.70, 0.73 Test: 0.71 and 0.74 respectively

Table 3: Thickness

MODEL	LEFT HEMISPHERE	RIGHT HEMISPHERE
K- Nearest Neighbors	We found the accuracy,F1 scores for Train: 0.98, 0.98 Test: 0.63 and 0.63 respectively	We found the accuracy,F1 scores for Train: 0.98, 0.98 Test: 0.57 and 0.58 respectively
Support Vector Machine	We found the accuracy,F1 scores for Train: 0.96, 0.96 Test: 0.67 and 0.66 respectively	We found the accuracy,F1 scores for Train: 0.94, 0.94 Test: 0.50 and 0.51 respectively
Multi-layer Perceptron (best)	We found the accuracy,F1 scores for Train: 0.81, 0.84 Test: 0.57 and 0.54 respectively	We found the accuracy,F1 scores for Train: 0.77, 0.80 Test: 0.53 and 0.55 respectively
Random Forest	We found the accuracy,F1 scores for Train: 1.0, 1.0 Test: 0.57 and 0.52 respectively	We found the accuracy,F1 scores for Train: 0.94, 0.94 Test: 0.67 and 0.66 respectively
Logistic Regression	We found the accuracy,F1 scores for Train: 0.74, 0.76 Test: 0.50 and 0.49 respectively	We found the accuracy,F1 scores for Train: 0.62, 0.65 Test: 0.53 and 0.54 respectively

While MRI studies in MDD have grown, the clinical impact of these studies has been small owing to the heterogeneity of both the findings and methodologies. By using ML models, we were able to overcome the reverse inference fallacy inherent to MRI methods. Modeling with machine learning not only automates the detection of depression but also offers the possibility of identifying biomarkers for it, which could further aid in understanding the correlation between brain morphometry and its functioning. A key objective of our study was to contribute to bridging the gap between depression research and the applied machine learning community, as ML techniques are increasingly used to analyze clinical studies.

The research we conducted is one of the few to combine three brain measurements together (cortical thickness, surface area, and regional volume) in a single study. This could potentially provide an alternative to traditional depression diagnosis techniques like behavioral observations and patient reports. Furthermore, this ML-based approach assists in understanding the progression of the disease by reporting the structural changes in the brain. Therefore, the process can be utilized as a prognostic tool in the future.

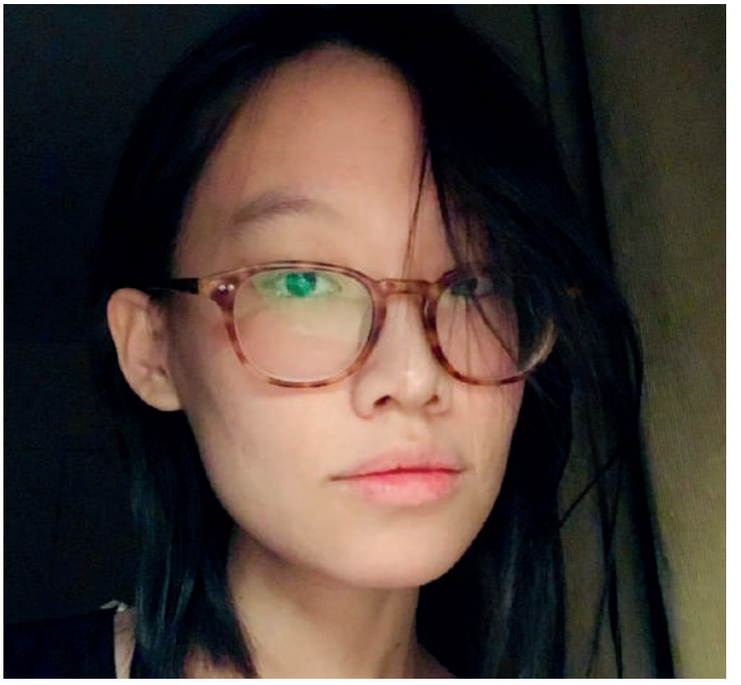
About Author:

SAI

Sai is a Visiting AI Researcher at deepkapha where she contributes to Neuro-AI projects. Her scientific interests lie in understanding the neural mechanisms underlying learning, brain plasticity and cognition focusing on perception, emotion, and consciousness. She wants to understand how emotion interacts with cognitive functions and draw insights to design better affective tools of mental health well-being.

Sai also holds a bachelors degree in Computer Science Engineering and her passion to study and reveal the mysteries of brain has led her to pursue a Neuroscience Masters at LMU Munich.

DR. JIE MEI



Dr. Jie Mei received her M.S. in cognitive neuroscience from École Normale Supérieure Paris in 2015, and completed her Ph.D. in medical neuroscience at Charité – Universitätsmediz in Berlin in 2019. Her major research interests include neuro-inspired AI, computational neuroscience, and application of machine learning in neuroscience and neurology. She is Head of AI Research at DeepKapha AI Research.

Spotlight Interview with

KEVIN TERRELL

CEO of BirchAI

BIO:

Kevin has over 20 years of experience as a go-to-market strategist and operational leader, with the bulk of that coming in healthcare. That experience includes leading marketing for an internal startup in the global healthcare practice at McKinsey & Company, and as the marketing and strategy leader for a division of General Electric, where he has also led process improvement for the commercial team. Kevin led business development in the healthcare sector at SparkCognition and started his career building products to exploit unstructured text at three-letter agencies in Washington, D.C.

Kevin earned his MBA at the University of Minnesota's Carlson School of Management, and his BA in German and Political Science at the University of Nebraska.

TELL US ABOUT YOUR JOURNEY IN AI. HOW HAS THE APPLICATION OF AI AND DATA SCIENCE EMPOWERED THE HEALTHCARE SECTOR?

I got my first exposure to AI in 2016 when I started doing business development for SparkCognition, where I focused on healthcare. I saw a big opportunity to automate workflows that rely on highly trained people working with unstructured language, but the technology was not quite ready for the tasks in healthcare. By 2020, when Yinhan, Sumant, and I started Birch, it was apparent that NLP had made the leap to where it could now start to automate tasks that required interpretation of complex, unstructured data.

WHAT IS BIRCHAI AND HOW DOES IT FIT INTO THE MODERN HEALTHCARE ECOSYSTEM?

Our initial focus at BirchAI is on solving the phone-related part of the equation. The problem for our customers being they have a small army of agents who are really good at talking to people and solving their problems, yet those agents spend up to 50% of their time documenting what they just did. At BirchAI, our first product helps solve that problem by automating complex After Call Work in a variety of healthcare call centers.

HOW DOES BIRCHAI USE ARTIFICIAL INTELLIGENCE FOR AUTOMATING HEALTHCARE CONTACT CENTER OPERATIONS? WHAT ARE SOME COMMON USE CASES OF BIRCHAI TECHNOLOGY?

We have a proprietary end-to-end pipeline that allows us to convert the phone conversation to text (Speech-to-Text), which feeds our NLP models that can classify and summarize those conversations within seconds. A good example would be any medical device firm with a patient who calls in to talk through an issue they are having with an implanted device. The person taking that call is highly trained, often someone who has a college degree in science. That agent spends seven or eight minutes troubleshooting the issue with the caller, and then spends about the same amount of time writing up and classifying that conversation. Our product can classify the call and write up the summary in less than 15 seconds – and you can't tell the difference between what people write and BirchAI. Similar use cases exist in the pharmaceutical industry, for payers (health insurers), and elsewhere in the industry.

HOW DOES MEDICALLY TRAINED, SPEECH-TO-TEXT TECHNOLOGY WORK? DOES TRANSFORMER-BASED NLP PLAY A ROLE IN BIRCHAI?

A big reason for that is that previous NLP technology had essentially plateaued in terms of its capabilities. But with transformer-based NLP we start at a higher level, with a lot of room to expand our capabilities, and by extension our solutions. It's really the combination of our own Speech-to-Text and NLP that makes the product work so well for customers. Of course, transformer-based NLP plays a role at BirchAI, and it will play a big role in the future of AI in healthcare.

WHAT ARE SOME OF THE BIGGEST CHALLENGES THAT BIRCHAI HAS HAD TO OVERCOME?

Labeling data is always a challenge for AI companies. But we have largely solved that with a complex AI-based pipeline to label data that we use to train at scale, and to reach a high degree of accuracy. Finding talented people who are a cultural fit is another challenge. We were able to hire a great founding engineer who is savvy on both software and large-scale AI. But finding employees like that who have a startup mentality is a lot of work.

HOW DOES BIRCHAI KEEP HEALTH RECORDS AND PATIENT DATA SECURE AND PRIVATE?

We have HIPAA-compliant reference architectures for both AWS and Azure that we use for our customers. Thus far customers have really appreciated how we approach data security, and that has enabled us to work with several very large healthcare companies.

WHAT DO YOU FORESEE AS SOME OF THE BIGGEST TRENDS AND CHALLENGES FOR AI AND NLP IN 2022 AND BEYOND?

AI adoption will accelerate as companies recognize they need to build capacity in the face of a declining workforce. For context – in the 30 years up to the 2008-era recession, the working population grew over 1.2% annually. That growth declined to about 0.50% until 2018 when Baby Boomers started to retire. Since that time the workforce has declined by about 0.20% per year, fed in part by the pandemic-induced “Great Resignation”. The healthcare industry is particularly threatened by this trend, as healthcare’s median employee age is five years older than the overall economy’s median employee age. Our customers see BirchAI as an effective way to fill that gap, while maintaining and even improving service levels.

Advancement of ARTIFICIAL INTELLIGENCE IN DENTISTRY

Artificial Intelligence (AI) is becoming more and more present in our lives nowadays. When you search on the internet, buying online, choosing where investing your money, tracing your route in a road trip, booking a hotel room, taking pictures in your smartphones and even when you go a medical doctor or a dentist, some steps of whole consultation may have hep from AI, such as: appointment, diagnosis support, imaging analysis and others.

Healthcare is one of the areas that most benefits from the help of AI. We've seen recently, during the COVID-19 Pandemic, the development of tools and systems that help pre-diagnosis the more critical sign of the disease – the Ground-Glass Opacity (GGO) in CT scans from the lungs and even in thorax X-rays. Another help of AI in this time was the record-breaking time of the Vaccine's development and availability worldwide.

Dental conditions such as caries, tooth loss, periodontitis and other mouth conditions afflict a large part of the global population with enormous cost for every country and dentistry can benefit from the use of AI in helping dentists in several aspects of the patient's treatment, for both quality and agility. Dentistry is divided in several specialties and most of the can take advantage of AI, the most common are:

- **Endodontics**
- **Implantology**
- **Oral and Maxillo-facial (OMF) Radiology**
- **Oral and Maxillo-facial (OMF) Surgery**
- **Orthodontics**
- **Periodontology**

ENDODONTICS

It is the root canal therapy (RCT), even if the tooth is vital (with a chronic or acute inflammation) or the tooth is non vital. The trained dental specialist, an endodontist, creates an access to the tooth pulp, removing it, cleaning, filling and sealing the root canal.

The AI can help the endodontists in diagnosing periapical lesions, root fractures or cracks, proximity of anatomical structures (the floor of the maxillary sinus, inferior dental canal), the working length (by the size of the root). There are recent studies of the AI helping in the prognosis of the RCT.

The most common dental imaging examination used during the RCT is periapical (intraoral), sometimes the panoramic of the face (OPG) for a broad analysis of both jaws and the Cone Beam Computed Tomography (CBCT) for a more accurate analysis of the tooth with RCT, once this is a 3D examination.

IMPLANTOLOGY

It is the specialty related to OMF Surgery where the dentist replaces a missing tooth, immediately after the extraction or not, with a dental implant made of titanium (most of the time). The base of dental implants is a biological process called osseointegration, where the titanium makes a bond to the adjacent bone; after a period of approximately 3 to 6 months, the prosthetics tooth crown can be attached to the dental implant.

The AI can help the implantology analyzing the proximity of anatomical structures (the floor of both maxillary sinus, inferior dental canal, mental foramen, incisive canal and foramen); the alveolar bone structure and measurements (hight, width and length) and the correct angulation of the alveolar bone for the dental implant related to the implant long axis.

The most common dental imaging examination used during the dental implant planning is the Cone Beam Computed Tomography (CBCT) for a more accurate analysis of jaws, due to its capability of showing any region in Multi Planar reconstruction (MPR), that are – Axial view – Coronal View – Sagittal View and the 3D reconstruction.

ORAL AND MAXILLO-FACIAL (OMF) SURGERY

It is the specialty related to the whole craniofacial complex (except the brain): mouth, both jaws, faces and the skull. It is his responsibility: mouth surgery (tooth extraction, impacted tooth extraction, etc); fixing fractures from both jaws; lesions and tumors removal in this area; reconstructive surgery in this area.

The AI can help OMF Surgery analyze the proximity of anatomical structures, the bone structure, predict diagnosis of some lesions, and virtual 3D surgery planning.

The most common dental imaging examination used during OMF Surgery is the Cone Beam Computed Tomography (CBCT) for a more accurate analysis of jaws, due to its capability of showing any region in Multi Planar reconstruction (MPR), that are – Axial view – Coronal View – Sagittal View and the 3D reconstruction; CT (Computed Tomography) Scans, depending on the region; in some cases, panoramic of the face (OPG) and periapical (intraoral) for a more specific and focal view.

ORTHODONTICS

It is the specialty that addresses the diagnosis, prevention, and correction of mal-positioned teeth and jaws, and misaligned bite patterns. It aims correcting the malocclusion and repositioning the dentition in relation to the craniofacial structures in the most harmonious way. The orthodontic treatment may require months even years to achieve the correct and expected result. The patient will have to use using dental braces and other appliances to gradually adjust tooth position and jaw alignment. In some severe cases, even jaw surgery may be required as part of the orthodontic treatment.

The AI can help Orthodontics evaluate the malocclusion problem, predict treatment planning, cephalometric analysis and tracing.

The most common dental imaging examination used during Orthodontics is the 2D Lateral Cephalometric X-ray helping localizing landmarks that are essential to create the cephalometric analyses and tracings; 3D Lateral Cephalometric from CBCT; panoramic X-ray of the face (OPG); facial and mouth photographs and study cast models.

PERIODONTOLOGY

It is the specialty that studies and treats the supporting and surrounding structures of the tooth: the gums (gingiva); the alveolar bone, the cementum and the periodontal ligament. The most common problems are gingivitis (some different levels); the presence of dental plaque; dental calculus. The most severe problem is periodontitis with 3 different categories: mild bone loss, moderate bone loss and severe bone loss. Some of these categories can be associated with a furcation lesion, that is the area between the

roots to form multi-roots teeth, such as molars and premolars. Depending on the severity of the periodontitis, the patient can have one or more teeth extracted.

Some risk factors may influence the severity of the periodontitis, such as: poor oral hygiene, not using dental floss, tobacco smoking, alcohol consumption and diabetes mellitus.

The AI can help Periodontology evaluate the periodontal diseases, including dental plaque, calculus, gingivitis, and periodontitis (bone level).

The most common dental imaging examination used during Periodontology is the; CBCT; panoramic X-ray of the face (OPG); periapical full mouth.

ORAL AND MAXILLO-FACIAL (OMF) RADIOLOGY

The OMF Radiology is the specialty concerned with the performance and the interpretation of any dental diagnostic imaging. There are several different types of dental imaging: panoramic (OPG), lateral cephalometric, frontal cephalometric, Temporomandibular joint (TMJ), periapical (intraoral), bite-wings (BW), occlusal superior and inferior, Cone Beam Computed Tomography (CBCT), hand and wrist.

In the last 15 years the OMF Radiology passed through a great development with the introduction of the concept of “filmless radiology” or Digital Radiology. With digital radiology there in no more x-ray film or chemical developing, so, the x-ray image is formed directly in the workstation computer (Direct Radiology) or a phosphor plate is used to acquire the latent x-ray image, then it is scanned with a proper device and formed in the workstation computer (Indirect Radiology). It was a great advance for radiology service, the referring dentist and for the patient. The service was favored with agility, zero waste (film and chemical), security and storage. The referring dentist receives the result faster, digitally and with better quality. The patient was the most benefited, because the digital radiology has a great decrease in radiation dose (in some cases, such as intraoral, up to 90% dose reduction), almost zero repetition, a faster and a digital result without the need to return to the service.

As seen above, in the other specialties, all of them use one or more types of dental diagnostic imaging, so the OMF Radiology is the most benefited with the help of AI. Most of the different types of dental imaging are used in Machine Learning (ML) models.

AI is been used in OMF Radiology for tooth identification, missing tooth, identify several radiographic findings, such as: decay(cavities), fillings (restorations), root canal therapy (RCT), periapical lesion, impacted tooth, dental implant planning, cephalometric landmark and analysis, dental implant identification, bone loss classification, identify anatomical structures and its proximity.

MACHINE LEARNING (ML)

One of the AI steps is the Machine Learning (ML), where the system will learn from data, identify patterns and make decisions with minimal human intervention with the use of a specific algorithm. When we talk about ML, there are, basically, 3 different types: Supervised Learning, Unsupervised Learning and Reinforcement Learning.

The most important and used type of ML in Dentistry is the Supervised Learning. The Supervised Learning is the task of learning a feature that maps an input to an output based on example input-output pairs, every single input data will receive a label that generates the output data. In dentistry for example, a panoramic x-ray (input data) where a human identifies each tooth (a label with the tooth identification) and the output data is the panoramic with every tooth identified.(Fig.01)

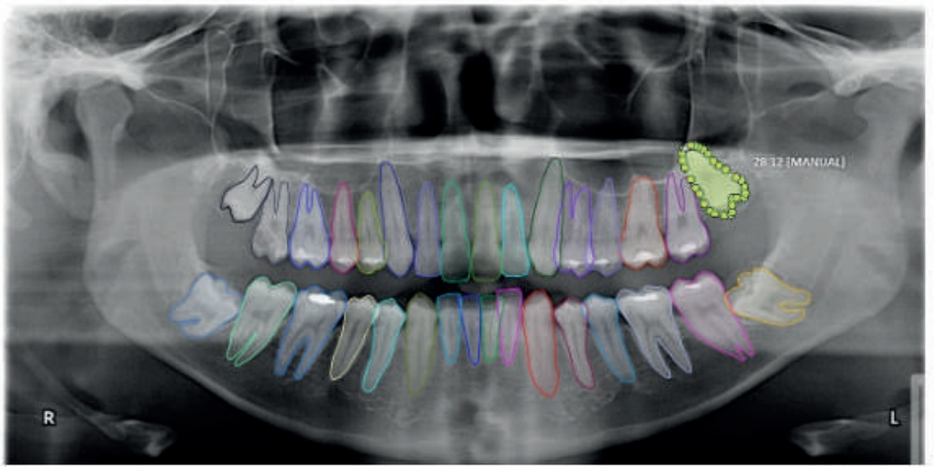


Fig.01 – Panoramic Annotated

THE MODEL

In Machine Learning, the model or dataset is the set of data that will be used in the learning process. The raw models are the original files (input data) that will go through the process of annotation (put the label in each file) and get the annotated dataset (output data). The model can be divided into Training Set, Validation Set and Test Set. The Training Set is the set of data that is used to train and make the model learn the hidden features/patterns in the data. The validation set is a set of data, separate from the training set, that is used to validate our model performance during training. The test set is a separate set of data used to test the model after completing the training.

ETHICS

One of the most important aspects of AI in healthcare, Medicine or Dentistry is Ethics. Once all models (datasets) are from a patient's images, the responsibility with data protection is huge!

Most of the countries have some laws or rules, in that regard. In the UE there is the GDPR (General Data Protection Regulation from April 27, 2016); in the USA there is the HIPAA (Health Insurance Portability and Accountability Act of 1996 from August 21, 1996; in Brazil there is the LGPD (General Law of Data Protection from August 16, 2020. What most of them have in common is that the patient data (name, sex, age, dob, etc) belongs to the patient and no one else, the data can’t be shared without the patient’s consent, the data must be removed from the database if the patient requires to.

DOCUMENTS AND STRATEGIES

With the advance of AI around the world, more and more countries are developing official documents to rule out the use of AI in every economic sector, including Healthcare. In a recent survey, a document produced by an organization in Rio de Janeiro – Brazil (https://itsrio.org/pt/home/), called Strategic Plan of AI Development from Mach, 2020 - (Planos Estratégicos de Desenvolvimento de Inteligência Artifical), identified the about 20 countries have already some document and 13 are in the process of developing one. In that survey, it was identified 5 international organization that has developed such a document. (Fig.02)

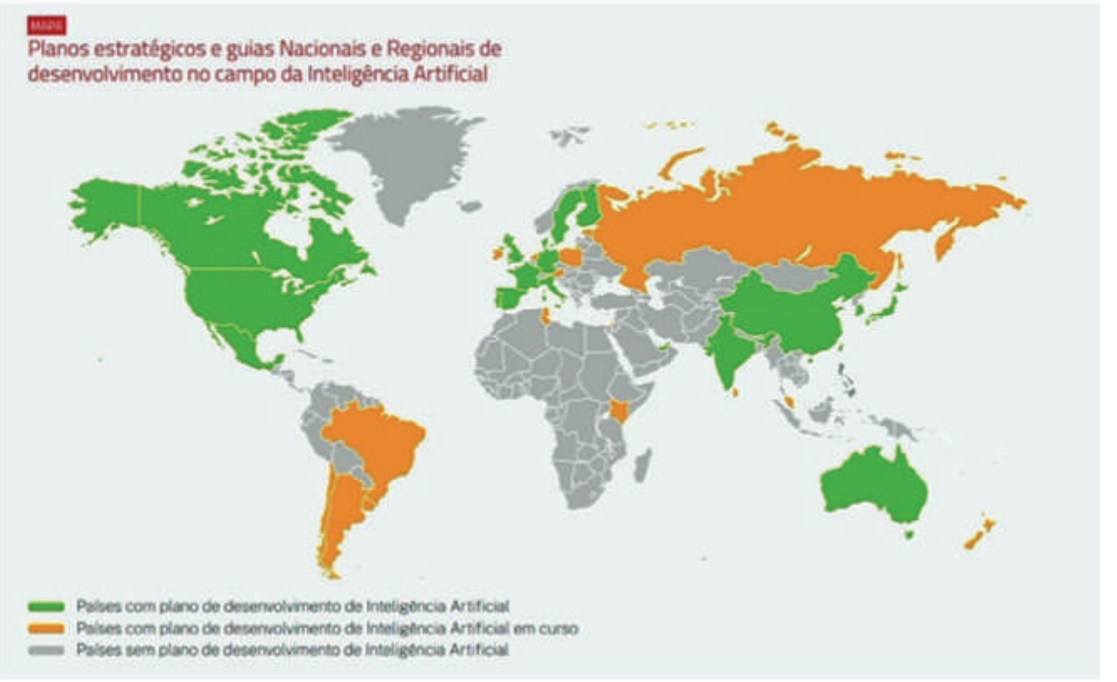


Fig.02 – World Map

COUNTRIES THAT HAVE A DOCUMENT

Australia	**Finland**
Japan	**Sweden**
Brazil	**France**
Mexico	**Taiwan**
Canada	**Germany**
Portugal	**United Arab Emirates**
China	**India**
Singapore	**United Kingdom**
Denmark	**Italy**
South Korea	**United States of America**

COUNTRIES THAT ARE DEVELOPING A DOCUMENT

Argentina
Chile
Ireland
Kenya
Malaysia
Netherland
New Zealand
Paraguay
Poland
Russia
Sri Lanka
Switzerland
Tunisia

INTERNATIONAL ORGANIZATION THAT HAS AN AI DOCUMENT

OECD – "Recommendation of the Council on Artificial Intelligence"

EU – Harmonizing Artificial Intelligence: "The role of standards in the EU AI regulation"

FDA - Artificial Intelligence / Machine Learning – Based Software as a Medical Device (SaMD) Action Plan

WHO - Ethics and Governance of Artificial Intelligence for Health WHO Guidance

IHE - Artificial Intelligence

- Workflow for Imaging"and Results
- Interoperability in Imaging (AIW-I)
- Results (AIR)

DR. NIELSEN SANTOS PEREIRA,DDS

OMF Radiology Certificate
Health Informatics Certificate
nielsen.spereira@gmail.com

I am from Rio de Janeiro, Brazil. I am 55 years old and I was graduated on Dentistry by the School of Dentistry of Federal University of Rio de Janeiro in 1987. I was certificate in Oral and Maxillofacial Radiology in 1991 by the same School of Dentistry. From 1988 until 1996 I served in the Brazilian Air Force as 1st Lieutenant Dentist. From 1992 to 2020 I was an entrepreneur and directed a group of dental clinics specialized in dental diagnostic imaging in Rio de Janeiro. In 2009 I developed a project for adopting digital radiology in all my clinics using the DICOM (Digital Imaging and Communication in Medicine), standard for interoperability and integrating all X-ray equipment with a Free Open Source PACS (Picture Archiving and Communication System). In 2010 I started working with CBCT (Cone Beam Computed Tomography), that is a MPR (Multi Planar Reconstruction) dental tomograph, highly used for dental implant planning among other indications. In 2013 I was certificate in Health Informatics by the Federal University of São Paulo. In the beginning of 2021, I started working on an Artificial Intelligence project, on my own, in dental diagnostic imaging, using Panoramic X-ray to identify teeth and some radiographic findings. I am currently collaborating with 'Deepkapha' in Netherl and to further accelerate the project.

Building Human Centered AI

deepkapha
AI Research

Bio:

Adityo Prakash is the co-founder and CEO of Verseon, a company at the forefront of modern drug discovery. Adityo started Verseon because he wanted to change how the world finds new medicines. He enjoys building fundamental science-based solutions to major business problems. In addition to guiding the company's business strategy, he has also been instrumental in building various aspects of the company's technology platform.

Prior to starting Verseon, Adityo was the co-founder and CEO of Pulsent, and the principal inventor of that company's technologies, which lie at the heart of all video streaming today. A technologist and business strategist with a track record of delivering industry firsts, he is an inventor on over 40 patent families, and received his BS in Mathematics and Physics from the California Institute of Technology.

Meet ADITYO PRAKASH (CEO of Verseon): A Technology-Based Pharmaceutical Company with a Proprietary Drug Discovery Platform

Tell us about Verseon. How does it fit into the modern pharmaceutical industry?

Verseon is a high-tech company whose end products are therapeutic small-molecule drugs that could change the standard of care for many of today's challenging diseases. Small-molecule drugs are the mainstay of the modern pharmacopeia. They are low-molecular weight organic molecules that can enter cells to interrupt disease processes and can typically be administered orally.

While pharmaceutical medicines—and small-molecule drugs in particular—have made great strides over the last century in treating diseases, many challenges remain. There are still many medical conditions that current medicines either cannot treat or can only treat poorly. Novel therapeutic interventions are critical to the future health and well-being of humanity.

At Verseon, our goal is to transform drug discovery and create a steady stream of precision treatments for major diseases. To

accomplish this, we've made a number of scientific breakthroughs in physics- and AI-driven molecule engineering that allow us to systematically design new, promising drugs that possess unique pharmacological properties and are unlikely to be found by any other current method. We have already successfully applied our platform to various diseases to generate a large number of highly promising drug candidates that are starting to progress through clinical trials.

What are some of the challenges in the drug discovery process that Verseon is trying to solve?

All other current methods of drug discovery rely on trial-and-error testing of a few million distinct compounds that have already been synthesized. And most of today's "AI-first" approaches focus on ingesting available experimental data and attempt to make predictions that might reduce the amount of trial-and-error testing within the field of known drug-like compounds. But these methods only skirt around the fundamental problem of drug discovery: the inability to find completely new drug molecules that have never before been synthesized anywhere. So, the industry typically finds "me-too" compounds that are very similar to previously studied molecules or repurposes existing drugs. And the ROI on drug research continues its long-term downward trend.

While the industry is currently limited to exploring a few million compounds and their associated data sets, a billion trillion trillion—or 1033—other possibilities remain completely unexplored. It's in this vast unknown ocean of possibilities that truly novel small molecules with unique pharmacological properties can be found.

That's where Verseon comes in.

We've spent a decade and a half developing fundamental advancements in computational chemistry, molecular-physics modeling, and applications of AI to make systematic design of completely new drug molecules a reality. For any disease-causing protein of interest, we can explore hundreds of millions of never-before synthesized drug-like molecular structures and computationally identify the best binders to advance into computer-directed synthesis and biological testing. Our proprietary physics modeling and AI-based methods allow us to sidestep the synthesis bottleneck and limited data sets that thwart the rest of the industry and find novel drug molecules with uniquely desirable properties that have the potential to change the standard of care for every disease we address.

How is AI helping in the drug discovery process for Verseon?

AI in various forms now permeates almost every step of our drug discovery process. But it is a thoughtful application of context-appropriate AI methodologies coupled with other scientific advances.

In some instances, AI helps improve the dynamic exploration of novel chemical matter by better leveraging knowledge associated with the target protein. Within our molecular-physics modeling engine, AI helps improve the accuracy and speed of our free-energy estimations. Other AI modules utilize comprehensive experimental test data from our novel chemicals to help us more efficiently explore adjacent regions of the chemical space and propose structural modifications that improve various pharmacological properties. Further downstream, there are early signs of positive benefit in utilizing AI-based analysis of clinical trials.

What are some of the diseases Verseon is initially targeting with its drug discovery platform? Can you name some of the novel drugs in your pipeline?

Most of our current drug programs fall under two broad categories: cardiometabolic disorders and cancers.

Within the cardiometabolic disorders category, I'll highlight two programs. Our Precision Oral Anticoagulants—or PROACs—and our oral prophylactics for diabetic vision loss.

Well over 100 million patients in the developed world need some form of life-long anticoagulant therapy to prevent heart attacks and strokes. Of those, a staggering 51 million patients could benefit from combination therapy with both an anticoagulant and an antiplatelet drug. Unfortunately, current anticoagulant drugs, many of them blockbusters from major pharma companies, all carry unacceptably high bleeding risk and are entirely unsuitable for long-term combined administration with an antiplatelet drug.

Our PROACs work through a unique mode of action by which they prevent clot formation and yet do not impair platelet function or significantly increase bleeding risk. This positions them to be the only drug candidates suitable for safe long-term combination therapy with antiplatelet drugs to prevent life-threatening strokes and heart attacks in this very large patient population.

Our most advanced PROAC candidate, code named VE-1902, has been delivering promising results and is nearing the end of Phase 1 trials. Our second candidate, VE-2851, has successfully completed preclinical tests and is ready to follow VE-1902 into clinical studies.

Diabetes is a disease that leads to many unpleasant life-altering problems. One of the most common is diabetic retinopathy—also known as DR—which affects over 35 million patients in the developed world. DR can lead to vision loss caused by blood vessel leakage in the retina. The current standard of care is to simply wait until vision loss is severe enough to justify regular injections into the eye with repurposed cancer drugs to treat the symptoms.

We have developed a novel class of oral drug candidates that could act as prophylactics, giving diabetics a chance to prevent this form of vision loss altogether. Our candidates are the only ones to successfully demonstrate reduction of blood vessel leakage in relevant in vivo tests. Our most advanced candidate, VE-4840, is ready to enter clinical trials.

In the arena of cancers, I'll mention two examples: our novel chemotherapy agents for multidrug resistant cancers and our first immuno-oncology program that targets an overexpressed protein present in half of all solid-tumor cancers.

While chemotherapy remains the first line of treatment for most cancers, too often resistance develops to multiple chemo drugs. Sadly, multidrug resistance causes 90% of cancer-related deaths that occur among chemo patients. In preclinical tests, Verseon's novel chemo drug candidates have shown that they don't lose effectiveness even when these cancers develop resistance to other drugs.

Our first immuno-oncology program focuses on CD73, a protein cancer cells frequently overexpress to mask their presence from the immune system. Verseon is developing a novel series of non-nucleotide CD73 inhibitors, fundamentally distinct from other compounds currently in development, that hold promise to reactivate immune response to CD73-positive cancers.

Again and again, what we find is that our ability to explore uncharted chemical space allows us to find drug candidates with uniquely desirable properties for the treatment of every disease we address.

Can you list some of the biggest challenges that Verseon has had to overcome?

Building a platform capable of systematic drug discovery and development is no small undertaking and required more than a decade of intensive work to achieve. As Sang Kim, our newly appointed CTO and a member of the National Academy of Engineering, so eloquently points out, "Verseon's platform is comprised of significant new advances within multiple distinct branches of science. Each of these advances would be enthusiastically welcomed by the leading practitioners in their respective domains—but the collection of these advances is virtually unattainable by any other organi-

zation."

Such a long development arc not only requires the right combination of talent, inspiration and luck to come together to solve hard scientific problems, but also the wherewithal and drive to weather various global macroeconomic events that can affect the entire industry.

Speaking of scientific and technological development, there were parts of the process we thought were sufficiently simple that we wouldn't have to build them in house and could just in-license. But when we needed to put those technologies in place, we discovered that no one else had ever built viable versions of those tools. In order to create a system that actually works, we had to build those pieces ourselves.

Navigating the major political and economic shifts during our platform development process was no mean feat either. For example, the 2008 crash impacted our development speed as economic activity and investment in our sector dramatically slowed. More recently, operational constraints posed by the current pandemic have included logistical challenges to conducting clinical trials, and companies like ourselves have had to adjust.

Despite various challenges over the past two decades, we have had a rewarding journey as we find ourselves in a position to positively impact global health.

What do you foresee as some of the biggest trends and challenges for AI in drug discovery?

In my view, the current state of AI in drug discovery strongly resembles the dot com boom of the late 1990s. Many companies simply use AI as a buzzword to make their products and services more marketable and bolster their share price. Such companies will see their prominence in the drug discovery space wane over time, and perhaps disappear when the inevitable shakeout comes. Other companies will create genuinely useful tools that will endure, but from what I've been able to observe, those companies will be comparatively few and far between.

My position is that the companies making the best use of AI will be the ones that couple AI techniques with significant advances in many other scientific arenas to make the discovery of completely new drugs more reliable, robust, and streamlined. Those companies will be the new leaders in the field of pharmaceutical medicine. Verseon plans to be foremost among them.

AI in HEALTHCARE

MARKTECHPOST

TO ADVERTISE ON MARKTECHPOST AI MAGAZINE
Contact :sales@marktechpost.com

Marktechpost, LLC.
1968 S. Coast Hwy, #1675
Laguna Beach, CA, 92651
USA

sales@marktechpost.com

www.marktechpost.com