Research Statement

Zongwei Zhou

I focus on early cancer detection. Almost everyone, whether connected personally or through friends or family members, has been affected by cancer. I envision a world where we can predict, well in advance, who is at high risk for cancer, enabling personalized screening and timely early detection far before the cancer has a chance to spread. This is a technically demanding challenge at the intersection of AI and medicine, and it motivates me to develop new AI algorithms in close collaboration with radiologists to ensure real clinical impact.

I lead the BodyMaps program within the Computational Cognition, Vision, and Learning (CCVL) research group at Johns Hopkins University (JHU). Our interdisciplinary group develops novel methodologies to improve early cancer detection by training computers to identify it in CT images and predict how it develops. This requires innovations that go beyond standard AI algorithms to ensure clinical relevance and real-world impact. Our work spans effective architectures, active learning, self- and weak supervision, large language and vision language models, image enhancement, and synthetic data generation. These contributions have been published and well recognized at top venues in AI/computer vision (CVPR, ICCV, NeurIPS, ICLR, TPAMI) and medical imaging (RSNA, MICCAI, AMIA, ISBI).

For example, at JHU, we have developed an AI algorithm that can detect and localize early pancreatic cancer from computed tomography images far more accurately than expert radiologists (sensitivity = 95% vs. 34%; specificity = 99%). This AI has been externally validated in 7,158 patients across North California, South California, Central Europe, Northwestern Europe, and Eastern Asia. This includes a national registry covering all pancreatic cancer diagnoses in Poland. *"This AI is very impressive,"* says a radiology specialist in Poland with extensive experience in imaging and treating pancreatic cancer. *"It catches the early precancerous lesions and small pancreatic cancers."*

Importantly, the AI can detect and localize cancer in the pancreas 13.6 months earlier than radiologists. If broadly deployed, it has the potential of improving five-year survival from 7-10% to 40-45% and reducing U.S. pancreatic cancer care costs by roughly $1 billion every year. The AI is now undergoing radiologist reader study and is being prepared for integration into clinical workflow. Our *retrospective* reader study shows that radiologists' sensitivity, with AI assistance, improves over 10% with a similar specificity. But future is the real test set. We are preparing for *prospective* study. The long-term goal is to obtain FDA approval and deploy this algorithm across all demographic groups worldwide—including underserved regions—to enable equitable early cancer detection. *"I was confused when radiologists and surgeons gave me completely different answers,"* recalls a JHU student who benefited from early detection and timely surgery. *"Pinning down the ground truth at early stages is difficult."*

Funded by multiple foundation and federal grants, including a recent NIH R01 (2025–2029), my group is extending these algorithms to additional cancer types. We currently detect cancers of the spleen, bladder, gallbladder, uterus, prostate, esophagus, and adrenal glands with performance that matches or exceeds reported radiologist accuracy on CT scans from Northern California and Eastern Asia. Across these organs, our models achieve 70% sensitivity and 88% specificity, outperforming radiologists by +15% and +5%, respectively. Moving forward, we want to detect a broad spectrum of cancers at earlier stages, classify them, and predict how they will develop over time. This motivates a unified, longitudinal AI that learns organ-specific and systemic changes to detect, classify, and forecast cancers years before symptoms.

This progress in early cancer detection builds on a decade of foundational AI research. My AI research is best known for creating UNet++, a segmentation architecture that has been cited more than 16,000 times since 2019. UNet++ redesigns the skip connections in U-Net, which makes it better at segmenting objects of different sizes—especially very small tumors that appear in early stages. It is now part of the standard toolkits in NVIDIA MONAI and Segmentation Models (PyTorch). It has also been applied outside medicine, such as in lunar crater detection, microseismic signal analysis, and image colorization, showing that its design works well whenever accurate segmentation is needed. I am also known for my work on self-supervised learning such as Models Genesis. *"The key contribution of this paper"*, says Professor Dinggang Shen, the General Chair of MICCAI-2019, *"is offering pre-trained models for 3D medical image analysis; it solves the previous constraint that only pre-trained 2D ImageNet models can be used for transfer learning in 3D medical imaging."* Moreover, I develop new AI algorithms—such as self-supervision, active learning, weak supervision, synthetic data generation, multi-modal 3D architectures, and spatially grounded vision–language models, as detailed below—that enable the early detection of subtle disease signals long before they become visible to clinicians.

In addition to developing AI that reliably detects early tumors, two major challenges remain: (1) assembling large, diverse datasets for rigorous multi-institutional evaluation, and (2) the prohibitive cost of voxel-wise tumor annotation for training, especially for early-stage tumors that are often missed in routine clinical practice. Unlike computer vision benchmarks such as ImageNet, medical images cannot be curated by crowdsourcing and annotated by non-experts. Having medical professionals annotate

images requires significant expertise, time, and efforts. These challenges motivate us to design new algorithms that can curate large-scale datasets and reduce the need for manual annotation.

To address the first challenge, we have assembled an unprecedented, multi-institutional resource. Unlike many AI studies that validate on single-site data, we draw on a de-identified, retrospective resource of more than 1,000,000 CT series from 155,000 patients across North America, Europe, and Asia (W. Li et al., NeurIPS 2025). These scans are paired, when available, with other imaging modalities (X-ray, mammography, MRI, ultrasound, PET), radiology and pathology reports, demographics, and risk factors from electronic health records (EHR). We created per-voxel tumor and organ annotations for key cohorts, enabling fine-grained learning of early tumors, as well as imaging biomarkers such as ductal dilation, parenchymal atrophy, and organ texture features. We pioneered the use of large language models (LLMs) to curate CT scans directly from the PACS and associated medical records (P. Bassi et al., ICCV 2025). This allows us to organize data for longitudinal studies and to assemble large-scale cohorts of pre-diagnostic and diagnostic cases. Because many patients have serial imaging, we can study cancer development over time by tracking biomarker trajectories, forecasting future changes, and inferring early disease onset. We are committed to open science and reproducible AI in medical imaging. To support the community, we have released image–mask–report datasets for over 10,000 patients from 145 hospitals worldwide (C. Qu et al., NeurIPS 2023; W. Li et al., MEDIA 2024; Q. Chen et al., ICCV 2025), catalyzing collaboration and facilitating translation of AI algorithms into clinical practice. These assets—rare in scale, diversity, and annotation depth—form the backbone of our research on early detection. Our interdisciplinary team includes experts in AI and medical imaging such as Prof. Alan Yuille (Computer Vision, Johns Hopkins University), Prof. Curtis Langlotz (Radiology, Stanford University), Prof. Kang Wang (Radiology, University of California, San Francisco), Prof. Heng Li (Radiation Oncology, Johns Hopkins Medicine), and Prof. Jarosław Ćwikła (Radiology, Poland), together with 12 pancreas specialists, 15 general radiologists, 23 radiology residents, and at least 50 medical trainees.

To address the second challenge, we have developed four approaches to substantially reduce the amount of manual annotation required. **First,** we advanced active learning workflows that make approximately 800×/80x more efficient for organ/tumor annotation than traditional annotation methods. The AI proposes candidate segmentations, check first by itself to create a priority list for radiologists, and radiologists simply accept, refine, or reject 5% of data with high priority. This approach allowed us to assemble large, high-quality annotated datasets and train reliable organ/tumor segmentation models (Y. Wang et al., MEDIA 2019; J. Liu et al., ICCV 2023), which in turn enable us to exploit global organ cues such as parenchymal atrophy and texture. **Second,** we developed a specialized vision–language model (VLM) that both extracts tumor information from radiology reports and generates clinically meaningful reports from images, outperforming standard VLMs (P. Bassi et al., ICCV 2025). By leveraging report-derived descriptions of tumor type, size, and approximate location, the model produces tens of thousands of weak labels across 14 organs, supporting large-scale multi-organ tumor detection. A dedicated training strategy converts these coarse report-derived annotations into effective supervision for tumor segmentation, improving tumor detection F1 scores by up to 16% (P. Bassi et al., MICCAI 2025, Best Paper Award runner-up). **Third,** we pioneered synthetic tumors as complementary training data. Realistic synthetic tumors increase sensitivity from 89% to 95% for small, subtle lesions that radiology reports often miss (Q. Hu et al., CVPR 2023) and generalize to other organs with visually similar early-stage tumors, such as liver and kidney (Q. Chen et al., CVPR 2024; ICCV 2025). They also accelerate active learning: generating a new tumor mask takes several minutes, whereas removing a false positive typically takes only a few seconds. This makes it efficient for the AI to "over-detect" and let radiologists prune false positives. **Finally,** we are developing image enhancement methods to improve tumor conspicuity in CT. These models translate non-contrast CT into high-quality virtual contrast phases (arterial, venous, delayed) and enhance tumor visibility on contrast-enhanced scans. This enables opportunistic screening from non-contrast CT, improving tumor detection F1 scores by roughly 10% (J. Liu et al., CVPR 2026 Under Review). We are also exploring cross-modal enhancement using PET and MRI to highlight tumors that appear more clearly in those modalities than on CT. These techniques have recently yielded exceptionally strong results for early-stage pancreatic tumors (W. Li et al., RSNA Oral 2025), with sensitivity/specificity approaching 95%/99% and a positive predictive value (PPV) of 84% and are being extended to other organs.

Over the next five years, my primary goal is to develop stronger AI algorithms that detect very small cancers (<2 cm) in the pancreas, hepatobiliary system, and upper-GI tract from two complementary CT sources: *indicated* scans (ordered for symptoms, abnormal labs, or known high risk) and opportunistic scans (acquired for other indications, such as trauma, often without contrast). For **indicated** scans, we aim for high sensitivity (>90%) and specificity (>95%) to minimize missed cancers and avoid unnecessary invasive follow-up. Our preliminary results are already strong: in high-risk populations, the AI attains 97% sensitivity and 99% specificity for pancreatic cancer. Because positive findings may prompt biopsy or surgery, PPV is crucial. When applied AI to high-risk cohorts—such as patients identified for blood tests—the pancreatic cancer prevalence rises to 4–5%. At this prevalence, the PPV of our existing AI algorithm exceeds 74%, meaning that most AI-flagged cases are true cancers and follow-up interventions are justified. We will further enhance performance in these indicated settings using hybrid supervision (voxel-wise labels, report-derived weak labels, and synthetic tumors) and imaging biomarkers such as ductal morphology and parenchymal texture to capture early causal changes in cancer development. For **opportunistic** scans, we analyze CTs from the general population obtained for unrelated reasons. This setting presents two key challenges: (1) we do not know who is high-

risk, and (2) most scans are non-contrast, making tumors hard for both radiologists and AI to detect. Because cancer prevalence is extremely low, our priority is to avoid missing early tumors, requiring very high sensitivity (>95%), while an acceptable specificity (>80%) is adequate given that follow-up imaging is low-risk. In such low-prevalence populations, PPV is naturally modest, so the goal is to capture early cues safely while ensuring that false positives lead only to non-invasive, short-interval follow-up. We will improve PPV through risk stratification using a medical knowledge graph and strengthen early-cue detection in non-contrast scans via domain adaptation (P. Bassi et al., Nature, In Preparation), contrast enhancement (J. Liu et al., CVPR 2026, Under Review), and longitudinal modeling.

To translate our AI algorithms to clinical settings, benefitting both radiologists and patients, and potentially for education purposes as well, we have already designed and conducted reader studies in pancreatic cancer for early detection (W. Li et al., Nature In Preparation) and report generation (P. Bassi et al., ICCV 2025). These studies will be extended to more cancer types. We will quantify reader-level improvements—such as higher sensitivity at matched specificity, reduced interpretation time with editable AI-drafted reports, and earlier detection measured by the interval between AI flag and clinical validation. We will also extend pre-specified subgroup analyses from the pancreas to other organs across demographics, scanner vendors, and imaging protocols (Q. Chen et al., CVPR 2026, Under Review). Building on these *retrospective* results, we will launch multi-institutional *prospective* trials, provided FDA approval, to determine whether earlier detection using our AI improves survival, reduces metastasis, and alters standard-of-care for high-risk populations. Our long-term goal is to conduct large-scale screening studies, including multicenter randomized trials comparing AI-assisted versus standard diagnostic pathways, to establish benefits in survival, morbidity, and healthcare utilization. We will measure actual patient outcomes—such as stage at detection, survival, cost savings, and reductions in invasive procedures—and partner with healthcare systems worldwide, including those in low- and middle-income countries, to ensure broad applicability and global health impact.

Beyond detection, I will use longitudinal data to uncover early tumor biology and pre-tumor changes—such as atrophy, ductal abnormalities, and texture shifts—that refine our understanding of tumorigenesis and may enable detection even earlier than current imaging thresholds. These insights could ultimately inform treatment guidelines for very small lesions, which remain poorly defined in clinical practice. This work requires close collaboration with pathologists, oncologists, radiologists, and AI experts to link imaging biomarkers with molecular alterations, genetic predisposition, and microenvironment changes. I will build on my prior contributions to develop the next generation of AI for early cancer detection. **First**, I will create multimodal 3D architectures that take advantage of all available information—not only CT images, but also radiology reports, imaging biomarkers, risk factors, and longitudinal history—to improve segmentation of subtle early-stage tumors. **Second**, I will develop human–AI annotation pipelines where strong models generate initial masks, EM-inspired self-labeling annotates most scans, and radiologists review only the most uncertain ~5% (W. Li et al., CVPR 2026, Under Review). Radiology reports will guide and validate these labels, ensuring clinically consistent tumor type, size, and location at scale. **Third**, I will develop spatially grounded vision–language models that understand 3D anatomy, organ–tumor–vessel relationships, and temporal changes. These models will align report text with precise 3D structures, reason about how tumors interact with surrounding anatomy over time, and predict future tumor development with clinically meaningful explanations. **Fourth**, I will develop generative AI models that simulate not only tumors but also their surrounding anatomy—ducts, vessels, and organs—and how they interact over time. These models will emulate tumor growth, metastasis, surgical scenarios, and treatment response, enabling data-driven prediction of disease progression and potentially informing updates to treatment plans and clinical guidelines. **Finally**, I will create image-enhancement algorithms that bring the strengths of multiple imaging modalities—such as the contrast detail of CT, soft-tissue characterization of MRI, and metabolic insight of PET—into a single, safer modality. These models will make tumors and other diseases more visible to both clinicians and AI, even from routine scans.

Ultimately, this research program aims to change how cancer is detected, diagnosed, and prevented. My goal is to shift discovery to much earlier and more treatable stages. I also seek to uncover the imaging and biological signatures that mark the earliest phases of tumor development. I hope to demonstrate that early detection using our algorithms can change the biology of disease progression—for example, by preventing metastatic spread or altering the natural history of cancer in a patient cohort.

In addition to my current funded NIH R01 project, I have seven NIH R01 proposals, one NIH R21 proposal, one NSF proposal, one DoD proposal, and several foundation and industry-sponsored proposals under review as Principal Investigator, totaling $30.95M in pending funding.