

WEBVTT

NOTE duration: "01:03:34.741"

NOTE Confidence: 0.9888475

00:00:00.399 --> 00:00:01.540 Good afternoon, everyone.

NOTE Confidence: 0.91186774

00:00:04.799 --> 00:00:06.080 It's a great pleasure today

NOTE Confidence: 0.91186774

00:00:06.080 --> 00:00:08.100 to host, doctor Faisal Mahboud,

NOTE Confidence: 0.9314906

00:00:09.280 --> 00:00:10.660 for the Department of Pathology

NOTE Confidence: 0.9314906

00:00:10.719 --> 00:00:11.460 Grand Rounds.

NOTE Confidence: 0.9496983

00:00:12.605 --> 00:00:13.805 He no he needs no

NOTE Confidence: 0.9496983

00:00:13.805 --> 00:00:14.925 introduction for most of us

NOTE Confidence: 0.9496983

00:00:14.925 --> 00:00:16.545 who work in anatomic pathology,

NOTE Confidence: 0.9496983

00:00:16.605 --> 00:00:17.885 who have watched with all

NOTE Confidence: 0.9496983

00:00:17.885 --> 00:00:18.385 the

NOTE Confidence: 0.9893424

00:00:18.845 --> 00:00:20.945 high impact and transform transformational

NOTE Confidence: 0.9893424

00:00:21.244 --> 00:00:22.204 work that he and his

NOTE Confidence: 0.9893424

00:00:22.204 --> 00:00:23.244 group have been engaged in

NOTE Confidence: 0.9893424

00:00:23.244 --> 00:00:24.465 in the last few years.

NOTE Confidence: 0.9389441

00:00:25.670 --> 00:00:27.770 Doctor. Mahmood is an engineer
NOTE Confidence: 0.9389441

00:00:27.830 --> 00:00:28.330 scientist
NOTE Confidence: 0.9050806

00:00:28.710 --> 00:00:30.810 after obtaining a bachelor's in
NOTE Confidence: 0.9529575

00:00:31.190 --> 00:00:33.030 electronic engineering from the GIK
NOTE Confidence: 0.9529575

00:00:33.030 --> 00:00:34.710 Institute of Engineering Sciences and
NOTE Confidence: 0.9529575

00:00:34.710 --> 00:00:35.930 Technology in Pakistan.
NOTE Confidence: 0.9844172

00:00:36.710 --> 00:00:37.590 He went on to get
NOTE Confidence: 0.9844172

00:00:37.590 --> 00:00:39.430 a PhD in biomedical engineering
NOTE Confidence: 0.9844172

00:00:39.430 --> 00:00:41.025 from the Okinawa Institute of
NOTE Confidence: 0.9844172

00:00:41.265 --> 00:00:42.245 Science and Technology.
NOTE Confidence: 0.9652578

00:00:43.345 --> 00:00:44.805 He then did his postdoctoral
NOTE Confidence: 0.9652578

00:00:44.945 --> 00:00:46.945 work in biomedical engineering at
NOTE Confidence: 0.9652578

00:00:46.945 --> 00:00:47.765 Johns Hopkins,
NOTE Confidence: 0.99959165

00:00:48.545 --> 00:00:50.225 where he started using his
NOTE Confidence: 0.99959165

00:00:50.225 --> 00:00:50.725 considerable
NOTE Confidence: 0.9955202

00:00:52.065 --> 00:00:53.905 expertise to address problems associated

NOTE Confidence: 0.9955202
00:00:53.905 --> 00:00:55.585 with using deep learning for
NOTE Confidence: 0.9955202
00:00:55.585 --> 00:00:56.485 medical imaging,
NOTE Confidence: 0.99568635
00:00:57.160 --> 00:00:59.340 particularly for endoscopy and pathology
NOTE Confidence: 0.99568635
00:00:59.400 --> 00:00:59.900 applications.
NOTE Confidence: 0.9947673
00:01:01.560 --> 00:01:02.700 He was then recruited,
NOTE Confidence: 0.95188
00:01:03.400 --> 00:01:04.760 to the Department of Pathology
NOTE Confidence: 0.95188
00:01:04.760 --> 00:01:06.360 at Harvard Medical School as
NOTE Confidence: 0.95188
00:01:06.360 --> 00:01:07.880 an assistant professor in two
NOTE Confidence: 0.95188
00:01:07.880 --> 00:01:08.620 thousand nineteen,
NOTE Confidence: 0.9985453
00:01:09.685 --> 00:01:11.365 where he has risen quickly
NOTE Confidence: 0.9985453
00:01:11.365 --> 00:01:12.584 to the rank of associate
NOTE Confidence: 0.9985453
00:01:12.645 --> 00:01:13.145 professor
NOTE Confidence: 0.8796117
00:01:13.605 --> 00:01:15.465 with appointments at Mass General,
NOTE Confidence: 0.8796117
00:01:15.685 --> 00:01:16.345 the Brigham,
NOTE Confidence: 0.9503735
00:01:16.805 --> 00:01:17.305 DFCI,
NOTE Confidence: 0.9859825

00:01:17.845 --> 00:01:19.065 and the Broad Institute.
NOTE Confidence: 0.98958576

00:01:20.725 --> 00:01:22.084 At Harvard, he now helms
NOTE Confidence: 0.98958576

00:01:22.084 --> 00:01:24.200 a large laboratory that's extensively
NOTE Confidence: 0.98958576

00:01:24.200 --> 00:01:26.120 funded by multiple agencies, including
NOTE Confidence: 0.98958576

00:01:26.120 --> 00:01:26.780 the NIH.
NOTE Confidence: 0.99455523

00:01:27.959 --> 00:01:29.159 The main focus of his
NOTE Confidence: 0.99455523

00:01:29.159 --> 00:01:31.159 lab is in developing AI
NOTE Confidence: 0.99455523

00:01:31.159 --> 00:01:33.020 tools for pathology image analysis.
NOTE Confidence: 0.9631953

00:01:34.280 --> 00:01:35.720 Doctor. Mahmood and his group
NOTE Confidence: 0.9631953

00:01:35.720 --> 00:01:37.095 have published extensively in this
NOTE Confidence: 0.9631953

00:01:37.095 --> 00:01:38.855 area with multiple high impact
NOTE Confidence: 0.9631953

00:01:38.855 --> 00:01:40.555 publications in top tier journals.
NOTE Confidence: 0.93686664

00:01:41.334 --> 00:01:42.775 Actually, by an informal count,
NOTE Confidence: 0.93686664

00:01:42.775 --> 00:01:43.735 I think in this past
NOTE Confidence: 0.93686664

00:01:43.735 --> 00:01:45.275 calendar year, he has
NOTE Confidence: 0.9470113

00:01:45.735 --> 00:01:48.055 one Nature, one Cell, three

NOTE Confidence: 0.9470113

00:01:48.055 --> 00:01:49.575 Nature Medicine papers, and that's

NOTE Confidence: 0.9470113

00:01:49.575 --> 00:01:50.850 just an informal count,

NOTE Confidence: 0.9554154

00:01:51.409 --> 00:01:52.450 and not mentioning some of

NOTE Confidence: 0.9554154

00:01:52.450 --> 00:01:53.110 the others.

NOTE Confidence: 0.96422386

00:01:54.210 --> 00:01:55.810 It's not surprising then that

NOTE Confidence: 0.96422386

00:01:55.810 --> 00:01:57.009 he has been the recipient

NOTE Confidence: 0.96422386

00:01:57.009 --> 00:01:59.250 of multiple awards, including the

NOTE Confidence: 0.96422386

00:01:59.250 --> 00:02:01.009 Outstanding Investigator Award from the

NOTE Confidence: 0.96422386

00:02:01.009 --> 00:02:02.229 NIH and NIGMS

NOTE Confidence: 0.9871829

00:02:03.810 --> 00:02:04.950 in twenty twenty.

NOTE Confidence: 0.95437336

00:02:05.815 --> 00:02:07.174 He's on the editorial boards

NOTE Confidence: 0.95437336

00:02:07.174 --> 00:02:08.875 of journals, including the AJP.

NOTE Confidence: 0.93032885

00:02:10.055 --> 00:02:11.415 He's the holder of several

NOTE Confidence: 0.93032885

00:02:11.415 --> 00:02:11.915 patents.

NOTE Confidence: 0.9831309

00:02:12.694 --> 00:02:13.815 And he has been a

NOTE Confidence: 0.9831309

00:02:13.815 --> 00:02:15.735 mentor to several graduate students,

NOTE Confidence: 0.9831309

00:02:15.735 --> 00:02:16.794 trainees, and postdocs.

NOTE Confidence: 0.918723

00:02:17.175 --> 00:02:18.375 And actually at last night

NOTE Confidence: 0.918723

00:02:18.375 --> 00:02:19.415 in dinner, he told us

NOTE Confidence: 0.918723

00:02:19.415 --> 00:02:21.080 the accomplishment that he's most

NOTE Confidence: 0.918723

00:02:21.080 --> 00:02:21.900 proud of

NOTE Confidence: 0.9796366

00:02:22.200 --> 00:02:23.560 is the fact that every

NOTE Confidence: 0.9796366

00:02:23.560 --> 00:02:24.760 postdoc he has trained has

NOTE Confidence: 0.9796366

00:02:24.760 --> 00:02:25.480 gone on to set up

NOTE Confidence: 0.9796366

00:02:25.480 --> 00:02:26.780 an independent lab.

NOTE Confidence: 0.9280188

00:02:27.400 --> 00:02:28.680 A testimony, I think, to

NOTE Confidence: 0.9280188

00:02:28.680 --> 00:02:30.599 his dedication to science and

NOTE Confidence: 0.9280188

00:02:30.599 --> 00:02:31.099 mentorship.

NOTE Confidence: 0.9555632

00:02:32.444 --> 00:02:33.905 We're very glad that he

NOTE Confidence: 0.9555632

00:02:34.044 --> 00:02:35.485 took some time to from

NOTE Confidence: 0.9555632

00:02:35.485 --> 00:02:37.084 his busy schedule to make

NOTE Confidence: 0.9555632
00:02:37.084 --> 00:02:38.305 the trip down from Boston
NOTE Confidence: 0.9662768
00:02:38.764 --> 00:02:39.965 to visit with us today
NOTE Confidence: 0.9662768
00:02:39.965 --> 00:02:40.465 and
NOTE Confidence: 0.9718784
00:02:41.165 --> 00:02:42.444 present and share some of
NOTE Confidence: 0.9718784
00:02:42.444 --> 00:02:43.245 the work that he and
NOTE Confidence: 0.9718784
00:02:43.245 --> 00:02:44.224 his group have
NOTE Confidence: 0.9971697
00:02:44.590 --> 00:02:45.650 been engaged in.
NOTE Confidence: 0.9435465
00:02:46.030 --> 00:02:47.069 The title of the talk
NOTE Confidence: 0.9435465
00:02:47.069 --> 00:02:47.969 is multimodal
NOTE Confidence: 0.9538131
00:02:48.430 --> 00:02:49.969 generative and agentic
NOTE Confidence: 0.7471835
00:02:50.430 --> 00:02:51.329 AI for pathology.
NOTE Confidence: 0.6953507
00:02:51.950 --> 00:02:52.769 Doctor. Mahdin.
NOTE Confidence: 0.87534106
00:02:54.269 --> 00:02:54.769 Okay.
NOTE Confidence: 0.9475298
00:02:58.035 --> 00:02:58.915 Thank you so much for
NOTE Confidence: 0.9475298
00:02:58.915 --> 00:03:00.355 the introduction and for inviting
NOTE Confidence: 0.9475298

00:03:00.355 --> 00:03:01.415 me to speak today.
NOTE Confidence: 0.9816608

00:03:02.115 --> 00:03:02.915 So I'll be talking a
NOTE Confidence: 0.9816608

00:03:02.915 --> 00:03:03.715 little bit about some of
NOTE Confidence: 0.9816608

00:03:03.715 --> 00:03:04.595 the work that my group
NOTE Confidence: 0.9816608

00:03:04.595 --> 00:03:05.555 has been doing over the
NOTE Confidence: 0.9816608

00:03:05.555 --> 00:03:07.315 past, six years or so
NOTE Confidence: 0.9816608

00:03:07.315 --> 00:03:08.935 in computational pathology.
NOTE Confidence: 0.9781652

00:03:09.555 --> 00:03:10.835 I'll talk about some of
NOTE Confidence: 0.9781652

00:03:10.835 --> 00:03:12.469 the older work that we
NOTE Confidence: 0.9781652

00:03:12.469 --> 00:03:12.969 did,
NOTE Confidence: 0.95800024

00:03:13.269 --> 00:03:14.069 and how it sort of
NOTE Confidence: 0.95800024

00:03:14.069 --> 00:03:14.950 led into some of the
NOTE Confidence: 0.95800024

00:03:14.950 --> 00:03:16.310 more recent work that's happened
NOTE Confidence: 0.95800024

00:03:16.310 --> 00:03:17.349 over the over the past
NOTE Confidence: 0.95800024

00:03:17.349 --> 00:03:18.870 two years. So there's just
NOTE Confidence: 0.95800024

00:03:18.870 --> 00:03:20.150 a quick outline for for

NOTE Confidence: 0.95800024
00:03:20.150 --> 00:03:20.950 the talk. I'll talk a
NOTE Confidence: 0.95800024
00:03:20.950 --> 00:03:22.310 little bit about weekly supervised
NOTE Confidence: 0.95800024
00:03:22.310 --> 00:03:23.769 learning for pathology,
NOTE Confidence: 0.9632777
00:03:24.794 --> 00:03:26.555 multimodal data integration and foundation
NOTE Confidence: 0.9632777
00:03:26.555 --> 00:03:28.794 models, generative AI, transitioning from
NOTE Confidence: 0.9632777
00:03:28.794 --> 00:03:29.595 two d to three d
NOTE Confidence: 0.9632777
00:03:29.595 --> 00:03:30.875 pathology, and bias and fairness
NOTE Confidence: 0.9632777
00:03:30.875 --> 00:03:33.115 in in, machine learning algorithms
NOTE Confidence: 0.9632777
00:03:33.115 --> 00:03:34.335 for pathology. So,
NOTE Confidence: 0.97513664
00:03:35.755 --> 00:03:37.840 a quick, note about about
NOTE Confidence: 0.97513664
00:03:37.840 --> 00:03:39.280 the problem formulation. So what
NOTE Confidence: 0.97513664
00:03:39.280 --> 00:03:40.080 what what are we trying
NOTE Confidence: 0.97513664
00:03:40.080 --> 00:03:41.040 to do? We're we're we're
NOTE Confidence: 0.97513664
00:03:41.040 --> 00:03:42.560 essentially trying to make sense
NOTE Confidence: 0.97513664
00:03:42.560 --> 00:03:43.360 of lots and lots of
NOTE Confidence: 0.97513664

00:03:43.360 --> 00:03:45.120 pathology data, images that look
NOTE Confidence: 0.97513664

00:03:45.120 --> 00:03:46.160 like this. I think everyone
NOTE Confidence: 0.97513664

00:03:46.160 --> 00:03:48.320 here is familiar with, with
NOTE Confidence: 0.97513664

00:03:48.320 --> 00:03:49.920 with with pathology images, but
NOTE Confidence: 0.97513664

00:03:49.920 --> 00:03:51.280 they're more like satellite images.
NOTE Confidence: 0.97513664

00:03:51.280 --> 00:03:52.965 They're they're hierarchical. They hold
NOTE Confidence: 0.97513664

00:03:52.965 --> 00:03:55.144 they hold information at multiple
NOTE Confidence: 0.9747572

00:03:55.444 --> 00:03:56.965 multiple levels. They're very different
NOTE Confidence: 0.9747572

00:03:56.965 --> 00:03:57.364 from,
NOTE Confidence: 0.9814463

00:03:58.165 --> 00:03:59.845 conventional computer vision kind of
NOTE Confidence: 0.9814463

00:03:59.845 --> 00:04:00.885 images that are used for
NOTE Confidence: 0.9814463

00:04:00.965 --> 00:04:02.165 to to develop machine learning
NOTE Confidence: 0.9814463

00:04:02.165 --> 00:04:03.765 algorithms. And the hope that
NOTE Confidence: 0.9814463

00:04:03.765 --> 00:04:04.805 we have is from just
NOTE Confidence: 0.9814463

00:04:04.805 --> 00:04:06.405 to go from large cohorts
NOTE Confidence: 0.9814463

00:04:06.405 --> 00:04:07.525 of these images to everything

NOTE Confidence: 0.9814463
00:04:07.525 --> 00:04:08.580 in the red box here.
NOTE Confidence: 0.9814463
00:04:08.580 --> 00:04:10.120 So early diagnosis, prognosis,
NOTE Confidence: 0.99930173
00:04:10.580 --> 00:04:11.640 prediction of response
NOTE Confidence: 0.8809692
00:04:11.940 --> 00:04:12.600 to treatment,
NOTE Confidence: 0.96684754
00:04:14.500 --> 00:04:16.260 integrated biomarker discovery, and so
NOTE Confidence: 0.96684754
00:04:16.260 --> 00:04:17.779 forth. So the
NOTE Confidence: 0.9713682
00:04:18.660 --> 00:04:19.779 what what we are essentially
NOTE Confidence: 0.9713682
00:04:19.779 --> 00:04:21.075 trying to do is train
NOTE Confidence: 0.9713682
00:04:21.075 --> 00:04:22.915 models using these images and
NOTE Confidence: 0.9713682
00:04:22.915 --> 00:04:24.535 labels that reside in pathology
NOTE Confidence: 0.9713682
00:04:24.595 --> 00:04:25.415 reports because
NOTE Confidence: 0.9795771
00:04:26.035 --> 00:04:27.635 they're the cheapest labels available.
NOTE Confidence: 0.9795771
00:04:27.635 --> 00:04:28.675 We we don't have lots
NOTE Confidence: 0.9795771
00:04:28.675 --> 00:04:30.035 and lots of, pixel level
NOTE Confidence: 0.9795771
00:04:30.035 --> 00:04:31.875 annotations for these for these
NOTE Confidence: 0.9795771

00:04:31.875 --> 00:04:32.375 images.
NOTE Confidence: 0.97682965

00:04:33.430 --> 00:04:34.390 So some of the earlier
NOTE Confidence: 0.97682965

00:04:34.390 --> 00:04:35.670 work we we did was
NOTE Confidence: 0.97682965

00:04:35.670 --> 00:04:37.670 around weekly supervised learning using
NOTE Confidence: 0.97682965

00:04:37.670 --> 00:04:38.570 whole slide images,
NOTE Confidence: 0.9567316

00:04:39.029 --> 00:04:40.790 slide level labels, and trying
NOTE Confidence: 0.9567316

00:04:40.790 --> 00:04:42.070 to make the make a
NOTE Confidence: 0.9567316

00:04:42.070 --> 00:04:43.830 conventional multiple instance learning based
NOTE Confidence: 0.9567316

00:04:43.830 --> 00:04:45.350 setup a little bit more
NOTE Confidence: 0.9567316

00:04:45.350 --> 00:04:46.630 more data efficient. So we
NOTE Confidence: 0.9567316

00:04:46.630 --> 00:04:48.275 used attention based multiple instance
NOTE Confidence: 0.9567316

00:04:48.275 --> 00:04:49.735 learning, pretrained encoders
NOTE Confidence: 0.9736933

00:04:50.195 --> 00:04:51.315 to see if if we
NOTE Confidence: 0.9736933

00:04:51.315 --> 00:04:52.835 can improve data efficiency. And
NOTE Confidence: 0.9736933

00:04:52.835 --> 00:04:54.354 this pipeline has been used
NOTE Confidence: 0.9736933

00:04:54.354 --> 00:04:54.854 extensively.

NOTE Confidence: 0.9816905
00:04:55.235 --> 00:04:56.835 So it's been used in
NOTE Confidence: 0.9816905
00:04:56.835 --> 00:04:58.135 over six hundred studies.
NOTE Confidence: 0.9685364
00:04:58.514 --> 00:04:59.670 We've seen it being used
NOTE Confidence: 0.9685364
00:04:59.750 --> 00:05:01.350 for with every major organ
NOTE Confidence: 0.9685364
00:05:01.350 --> 00:05:02.630 organ system all the way
NOTE Confidence: 0.9685364
00:05:02.630 --> 00:05:03.750 up to forensics. So it's
NOTE Confidence: 0.9685364
00:05:03.750 --> 00:05:05.029 very exciting to see how
NOTE Confidence: 0.9685364
00:05:05.029 --> 00:05:06.470 people are beginning to to
NOTE Confidence: 0.9685364
00:05:06.470 --> 00:05:08.010 to use this. But,
NOTE Confidence: 0.9938358
00:05:09.110 --> 00:05:10.730 a few applications that we,
NOTE Confidence: 0.9824404
00:05:11.190 --> 00:05:12.650 we we sort of
NOTE Confidence: 0.952371
00:05:13.574 --> 00:05:15.095 targeted when when we first
NOTE Confidence: 0.952371
00:05:15.095 --> 00:05:16.214 developed this, the first one
NOTE Confidence: 0.952371
00:05:16.214 --> 00:05:17.255 was for cancer defined in
NOTE Confidence: 0.952371
00:05:17.255 --> 00:05:18.615 primary where we tried to
NOTE Confidence: 0.952371

00:05:18.615 --> 00:05:20.294 show that conventional h and
NOTE Confidence: 0.952371

00:05:20.294 --> 00:05:21.414 e images could be used
NOTE Confidence: 0.952371

00:05:21.414 --> 00:05:22.794 to predict what the origin
NOTE Confidence: 0.952371

00:05:22.935 --> 00:05:23.895 of the tumor may be
NOTE Confidence: 0.952371

00:05:23.895 --> 00:05:25.255 or give indications for what
NOTE Confidence: 0.952371

00:05:25.255 --> 00:05:27.230 would be the top top
NOTE Confidence: 0.952371

00:05:27.230 --> 00:05:29.070 three, top five predictions. And
NOTE Confidence: 0.952371

00:05:29.070 --> 00:05:30.430 then those predictions could be
NOTE Confidence: 0.952371

00:05:30.430 --> 00:05:31.630 could be used to and
NOTE Confidence: 0.952371

00:05:31.630 --> 00:05:32.830 then the origin could be
NOTE Confidence: 0.952371

00:05:32.830 --> 00:05:35.070 confirmed with additional ancillary tests.
NOTE Confidence: 0.952371

00:05:35.070 --> 00:05:35.790 That that's what led to
NOTE Confidence: 0.952371

00:05:35.790 --> 00:05:36.589 this study, and we did
NOTE Confidence: 0.952371

00:05:36.589 --> 00:05:37.950 a lot of analysis on
NOTE Confidence: 0.952371

00:05:37.950 --> 00:05:40.110 internal cohorts, external cohorts. This
NOTE Confidence: 0.952371

00:05:40.110 --> 00:05:41.710 really confirmed that it was

NOTE Confidence: 0.952371
00:05:41.710 --> 00:05:43.725 possible to predict origins directly
NOTE Confidence: 0.952371
00:05:43.725 --> 00:05:44.385 from histology.
NOTE Confidence: 0.9957699
00:05:45.725 --> 00:05:46.225 Before
NOTE Confidence: 0.9632651
00:05:46.525 --> 00:05:47.645 this work was done, there
NOTE Confidence: 0.9632651
00:05:47.645 --> 00:05:48.525 are a number of different
NOTE Confidence: 0.9632651
00:05:48.525 --> 00:05:49.645 studies showing that you could
NOTE Confidence: 0.9632651
00:05:49.645 --> 00:05:50.145 predict,
NOTE Confidence: 0.9861675
00:05:50.685 --> 00:05:52.705 origins from lots of molecular
NOTE Confidence: 0.9861675
00:05:52.765 --> 00:05:53.265 data,
NOTE Confidence: 0.983667
00:05:53.965 --> 00:05:55.405 and studies showing that you
NOTE Confidence: 0.983667
00:05:55.405 --> 00:05:57.325 could predict molecular alterations directly
NOTE Confidence: 0.983667
00:05:57.325 --> 00:05:58.365 from histology. So if the
NOTE Confidence: 0.983667
00:05:58.365 --> 00:05:59.940 two statements are true, it
NOTE Confidence: 0.983667
00:05:59.940 --> 00:06:00.979 should be possible to predict
NOTE Confidence: 0.983667
00:06:00.979 --> 00:06:02.360 origins directly from,
NOTE Confidence: 0.97180486

00:06:02.819 --> 00:06:04.660 from from from metastatic images.
NOTE Confidence: 0.97180486

00:06:04.660 --> 00:06:05.460 So that's what we showed
NOTE Confidence: 0.97180486

00:06:05.460 --> 00:06:06.279 in this study.
NOTE Confidence: 0.9770224

00:06:07.460 --> 00:06:08.500 More recently, what we have
NOTE Confidence: 0.9770224

00:06:08.500 --> 00:06:09.460 done is that we've tried
NOTE Confidence: 0.9770224

00:06:09.460 --> 00:06:11.300 to expand this into a
NOTE Confidence: 0.9770224

00:06:11.300 --> 00:06:11.800 multimodal
NOTE Confidence: 0.9064235

00:06:12.435 --> 00:06:14.115 integration based analysis where we're
NOTE Confidence: 0.9064235

00:06:14.115 --> 00:06:15.335 trying to integrate,
NOTE Confidence: 0.94605094

00:06:16.355 --> 00:06:18.695 histology images with molecular data
NOTE Confidence: 0.94605094

00:06:18.835 --> 00:06:19.555 to see if we can
NOTE Confidence: 0.94605094

00:06:19.555 --> 00:06:20.675 if we can improve origin
NOTE Confidence: 0.94605094

00:06:20.675 --> 00:06:21.875 prediction, and we showed that
NOTE Confidence: 0.94605094

00:06:21.875 --> 00:06:23.495 that was that was possible.
NOTE Confidence: 0.99341315

00:06:24.035 --> 00:06:24.915 And then just speaking of
NOTE Confidence: 0.99341315

00:06:24.915 --> 00:06:26.675 multimodal, we're also interested in

NOTE Confidence: 0.99341315

00:06:26.675 --> 00:06:28.780 combining histology images with with

NOTE Confidence: 0.99341315

00:06:28.780 --> 00:06:30.380 molecular data to see if

NOTE Confidence: 0.99341315

00:06:30.380 --> 00:06:31.200 we can combine,

NOTE Confidence: 0.995842

00:06:31.740 --> 00:06:33.280 combine the two to improve

NOTE Confidence: 0.995842

00:06:33.420 --> 00:06:34.480 outcomes. So,

NOTE Confidence: 0.957078

00:06:35.020 --> 00:06:36.620 it's been shown extensively that

NOTE Confidence: 0.957078

00:06:36.620 --> 00:06:37.900 you can use histology whole

NOTE Confidence: 0.957078

00:06:37.900 --> 00:06:38.640 slide images

NOTE Confidence: 0.8916134

00:06:39.100 --> 00:06:39.500 and,

NOTE Confidence: 0.9406205

00:06:39.900 --> 00:06:41.660 and separate patients into distinct

NOTE Confidence: 0.9406205

00:06:41.660 --> 00:06:42.675 distinct risk groups,

NOTE Confidence: 0.94121623

00:06:43.395 --> 00:06:45.154 using weekly supervised algorithms. And

NOTE Confidence: 0.94121623

00:06:45.154 --> 00:06:45.875 you could do the same

NOTE Confidence: 0.94121623

00:06:45.875 --> 00:06:47.875 with, with molecular data, whether

NOTE Confidence: 0.94121623

00:06:47.875 --> 00:06:49.395 it's NGS or a combination

NOTE Confidence: 0.94121623

00:06:49.395 --> 00:06:51.254 of NGS and transcriptomic data.
NOTE Confidence: 0.95986384

00:06:51.714 --> 00:06:52.675 In this study, we wanted
NOTE Confidence: 0.95986384

00:06:52.675 --> 00:06:53.235 to see if we could
NOTE Confidence: 0.95986384

00:06:53.315 --> 00:06:54.294 if you could do this
NOTE Confidence: 0.9801314

00:06:55.930 --> 00:06:57.370 exhaustively in a pan cancer
NOTE Confidence: 0.9801314

00:06:57.370 --> 00:06:57.870 manner,
NOTE Confidence: 0.97674334

00:06:58.970 --> 00:07:00.010 for lots of different cancer
NOTE Confidence: 0.97674334

00:07:00.010 --> 00:07:00.970 types. And we showed that
NOTE Confidence: 0.97674334

00:07:00.970 --> 00:07:02.170 you can separate patients into
NOTE Confidence: 0.97674334

00:07:02.170 --> 00:07:03.529 into distinct risk groups and
NOTE Confidence: 0.97674334

00:07:03.529 --> 00:07:05.610 then improve patient stratification by
NOTE Confidence: 0.97674334

00:07:05.610 --> 00:07:07.450 integrating additional modalities. But then
NOTE Confidence: 0.97674334

00:07:07.450 --> 00:07:08.169 you can also go back
NOTE Confidence: 0.97674334

00:07:08.169 --> 00:07:08.970 and look at what was
NOTE Confidence: 0.97674334

00:07:08.970 --> 00:07:10.510 important in the molecular profile,
NOTE Confidence: 0.97674334

00:07:10.685 --> 00:07:12.045 what's important in the in

NOTE Confidence: 0.97674334
00:07:12.045 --> 00:07:13.665 the histology, and further quantify
NOTE Confidence: 0.97674334
00:07:13.725 --> 00:07:14.865 what the model essentially
NOTE Confidence: 0.96034414
00:07:15.245 --> 00:07:16.605 pays attention to and using
NOTE Confidence: 0.96034414
00:07:16.605 --> 00:07:18.365 interpretability as a as a
NOTE Confidence: 0.96034414
00:07:18.365 --> 00:07:19.805 discovery mechanism. So there are
NOTE Confidence: 0.96034414
00:07:19.805 --> 00:07:21.165 lots of lots of,
NOTE Confidence: 0.996151
00:07:21.965 --> 00:07:23.725 interesting findings in this, in
NOTE Confidence: 0.996151
00:07:23.725 --> 00:07:24.465 this study.
NOTE Confidence: 0.92464113
00:07:26.480 --> 00:07:27.380 We also
NOTE Confidence: 0.9777138
00:07:27.760 --> 00:07:29.300 did some work on
NOTE Confidence: 0.9178133
00:07:30.240 --> 00:07:32.400 endomyocardial biopsy assessment trying to
NOTE Confidence: 0.9178133
00:07:32.400 --> 00:07:34.100 see if we can improve,
NOTE Confidence: 0.9300463
00:07:34.960 --> 00:07:35.460 standardization
NOTE Confidence: 0.858623
00:07:35.920 --> 00:07:36.420 for,
NOTE Confidence: 0.96636933
00:07:38.320 --> 00:07:39.540 for cardiac biopsies
NOTE Confidence: 0.8975183

00:07:39.840 --> 00:07:41.220 after after heart transplants.

NOTE Confidence: 0.9399133

00:07:42.325 --> 00:07:43.525 After patients get get a

NOTE Confidence: 0.9399133

00:07:43.525 --> 00:07:44.565 heart transplant, they would often

NOTE Confidence: 0.9399133

00:07:44.565 --> 00:07:46.805 get repeated endomyocardial biopsies to

NOTE Confidence: 0.9399133

00:07:46.805 --> 00:07:48.085 see if the donor heart

NOTE Confidence: 0.9399133

00:07:48.085 --> 00:07:49.205 is being rejected by the

NOTE Confidence: 0.9399133

00:07:49.205 --> 00:07:50.245 by the recipient and it's

NOTE Confidence: 0.9399133

00:07:50.245 --> 00:07:51.205 a problem where there's a

NOTE Confidence: 0.9399133

00:07:51.205 --> 00:07:52.645 large scale intra and intra

NOTE Confidence: 0.9399133

00:07:52.645 --> 00:07:54.425 observer variability. So this study,

NOTE Confidence: 0.9399133

00:07:54.645 --> 00:07:56.185 we try to focus on,

NOTE Confidence: 0.99752903

00:07:56.850 --> 00:07:58.930 just assessing whether human in

NOTE Confidence: 0.99752903

00:07:58.930 --> 00:08:00.710 the loop AI can improve

NOTE Confidence: 0.98726803

00:08:01.169 --> 00:08:01.669 standardization.

NOTE Confidence: 0.9587676

00:08:02.130 --> 00:08:03.330 So I'll refer you to

NOTE Confidence: 0.9587676

00:08:03.330 --> 00:08:04.790 the article for for more

NOTE Confidence: 0.9587676

00:08:05.010 --> 00:08:07.190 more in-depth analysis on this.

NOTE Confidence: 0.9587676

00:08:07.330 --> 00:08:08.530 But the idea of showing

NOTE Confidence: 0.9587676

00:08:08.530 --> 00:08:09.729 all of these examples is

NOTE Confidence: 0.9587676

00:08:09.729 --> 00:08:11.110 that most of the

NOTE Confidence: 0.9426628

00:08:11.755 --> 00:08:13.595 studies in computational pathology follow

NOTE Confidence: 0.9426628

00:08:13.595 --> 00:08:14.635 some form of this this

NOTE Confidence: 0.9426628

00:08:14.635 --> 00:08:15.935 framework where you have,

NOTE Confidence: 0.9450426

00:08:17.515 --> 00:08:18.955 images, pathology slides that are

NOTE Confidence: 0.9450426

00:08:18.955 --> 00:08:20.575 digitized into whole slide images.

NOTE Confidence: 0.9450426

00:08:20.715 --> 00:08:21.755 And the whole slide images

NOTE Confidence: 0.9450426

00:08:21.755 --> 00:08:24.095 get processed by segmentation, patching.

NOTE Confidence: 0.9450426

00:08:24.235 --> 00:08:25.080 And then you have some

NOTE Confidence: 0.9450426

00:08:25.080 --> 00:08:26.520 form of feature extraction, feature

NOTE Confidence: 0.9450426

00:08:26.520 --> 00:08:28.599 aggregation and prediction. Right? So

NOTE Confidence: 0.9450426

00:08:28.599 --> 00:08:30.120 a majority of computational topology

NOTE Confidence: 0.9450426

00:08:30.120 --> 00:08:31.580 studies have followed this pipeline.
NOTE Confidence: 0.98826426

00:08:31.960 --> 00:08:33.480 On the technical end, most
NOTE Confidence: 0.98826426

00:08:33.480 --> 00:08:34.839 improvements are made either in
NOTE Confidence: 0.98826426

00:08:34.839 --> 00:08:36.459 feature extraction or feature aggregation.
NOTE Confidence: 0.9483519

00:08:37.395 --> 00:08:38.275 Over the past two years,
NOTE Confidence: 0.9483519

00:08:38.275 --> 00:08:39.975 we realized that feature extraction
NOTE Confidence: 0.9483519

00:08:40.035 --> 00:08:41.335 is way more important
NOTE Confidence: 0.96687984

00:08:41.795 --> 00:08:43.315 in comparison to feature aggregation.
NOTE Confidence: 0.96687984

00:08:43.315 --> 00:08:44.535 If you have richer features,
NOTE Confidence: 0.96687984

00:08:44.675 --> 00:08:45.655 you you can,
NOTE Confidence: 0.9954325

00:08:46.995 --> 00:08:48.535 get to a better prediction.
NOTE Confidence: 0.8700119

00:08:49.715 --> 00:08:51.149 And the feature extraction is
NOTE Confidence: 0.8700119

00:08:51.149 --> 00:08:52.350 sort of what's driven the
NOTE Confidence: 0.8700119

00:08:52.350 --> 00:08:54.110 foundation model evolution pathology and
NOTE Confidence: 0.8700119

00:08:54.110 --> 00:08:55.309 foundation model are sort of
NOTE Confidence: 0.8700119

00:08:55.309 --> 00:08:55.809 or

NOTE Confidence: 0.8712631
00:08:56.190 --> 00:08:56.910 or or hype, but all
NOTE Confidence: 0.8712631
00:08:56.910 --> 00:08:58.110 they are are they're large
NOTE Confidence: 0.8712631
00:08:58.110 --> 00:08:59.410 self supervised models.
NOTE Confidence: 0.8123131
00:09:00.190 --> 00:09:02.130 They extract the feature representations
NOTE Confidence: 0.96759087
00:09:02.670 --> 00:09:03.790 from these images. And if
NOTE Confidence: 0.96759087
00:09:03.790 --> 00:09:05.230 you have retrofit representations, you
NOTE Confidence: 0.96759087
00:09:05.230 --> 00:09:06.225 hope that you can use
NOTE Confidence: 0.96759087
00:09:06.464 --> 00:09:07.904 fewer data points for downstream
NOTE Confidence: 0.96759087
00:09:07.904 --> 00:09:08.404 training,
NOTE Confidence: 0.9965363
00:09:09.985 --> 00:09:10.865 and these models can be
NOTE Confidence: 0.9965363
00:09:10.865 --> 00:09:12.704 applicable to rare diseases, clinical
NOTE Confidence: 0.9965363
00:09:12.704 --> 00:09:13.204 trials,
NOTE Confidence: 0.98746616
00:09:13.904 --> 00:09:15.184 or situations where you just
NOTE Confidence: 0.98746616
00:09:15.184 --> 00:09:16.545 have fewer fewer data points
NOTE Confidence: 0.98746616
00:09:16.545 --> 00:09:18.144 available. Foundation models are not
NOTE Confidence: 0.98746616

00:09:18.144 --> 00:09:19.345 really meant to replace self
NOTE Confidence: 0.98746616

00:09:19.345 --> 00:09:21.285 supervised task specific models.
NOTE Confidence: 0.93703693

00:09:21.650 --> 00:09:23.170 They aid with the development
NOTE Confidence: 0.93703693

00:09:23.170 --> 00:09:25.030 of better task specific
NOTE Confidence: 0.96652496

00:09:26.610 --> 00:09:28.850 models. And the the the
NOTE Confidence: 0.96652496

00:09:28.850 --> 00:09:30.710 two foundation models from the
NOTE Confidence: 0.90408534

00:09:31.170 --> 00:09:32.370 from the group, that came
NOTE Confidence: 0.90408534

00:09:32.370 --> 00:09:33.910 out in March this year,
NOTE Confidence: 0.90408534

00:09:34.005 --> 00:09:34.804 We call them Unni and
NOTE Confidence: 0.90408534

00:09:34.804 --> 00:09:35.845 Panch. It was an image
NOTE Confidence: 0.90408534

00:09:35.845 --> 00:09:37.605 based model which uses lots
NOTE Confidence: 0.90408534

00:09:37.605 --> 00:09:38.745 of pathology images,
NOTE Confidence: 0.965945

00:09:39.365 --> 00:09:40.565 in a self supervised manner
NOTE Confidence: 0.965945

00:09:40.565 --> 00:09:41.684 and an image text model
NOTE Confidence: 0.965945

00:09:41.684 --> 00:09:42.345 that contrasts,
NOTE Confidence: 0.99044114

00:09:43.365 --> 00:09:45.225 images with text to improve

NOTE Confidence: 0.972149
00:09:45.525 --> 00:09:47.465 essentially image based feature representation.
NOTE Confidence: 0.9548962
00:09:48.165 --> 00:09:49.179 So for the image based
NOTE Confidence: 0.9548962
00:09:49.179 --> 00:09:50.059 model, we use a hundred
NOTE Confidence: 0.9548962
00:09:50.059 --> 00:09:51.580 thousand whole slide images, a
NOTE Confidence: 0.9548962
00:09:51.580 --> 00:09:53.179 hundred million patches from these
NOTE Confidence: 0.9548962
00:09:53.179 --> 00:09:54.700 hundred, hundred thousand whole slide
NOTE Confidence: 0.9548962
00:09:54.700 --> 00:09:56.320 images is just a distribution
NOTE Confidence: 0.9548962
00:09:56.540 --> 00:09:57.980 for where all the data
NOTE Confidence: 0.9548962
00:09:57.980 --> 00:09:59.260 came from. We try to
NOTE Confidence: 0.9548962
00:09:59.260 --> 00:10:01.339 maximize for diversity. So trying
NOTE Confidence: 0.9548962
00:10:01.339 --> 00:10:02.640 to deliberately collect
NOTE Confidence: 0.89677095
00:10:03.005 --> 00:10:03.965 data that,
NOTE Confidence: 0.98285306
00:10:05.565 --> 00:10:07.005 maximize the diversity we had
NOTE Confidence: 0.98285306
00:10:07.005 --> 00:10:08.705 in the overall overall dataset.
NOTE Confidence: 0.984679
00:10:10.205 --> 00:10:11.485 And we did not use
NOTE Confidence: 0.984679

00:10:11.485 --> 00:10:12.845 any public data in in
NOTE Confidence: 0.984679

00:10:12.845 --> 00:10:13.885 developing this model. And that
NOTE Confidence: 0.984679

00:10:13.885 --> 00:10:15.005 was on purpose because we
NOTE Confidence: 0.984679

00:10:15.005 --> 00:10:16.320 wanted to use lots of
NOTE Confidence: 0.984679

00:10:16.400 --> 00:10:18.420 public data as as independent
NOTE Confidence: 0.984679

00:10:18.559 --> 00:10:19.860 evaluation cohorts.
NOTE Confidence: 0.97096187

00:10:21.120 --> 00:10:22.320 It was trained using the
NOTE Confidence: 0.97096187

00:10:22.320 --> 00:10:23.840 Dyno Dyno v two framework,
NOTE Confidence: 0.97096187

00:10:23.840 --> 00:10:25.140 and it was compared against
NOTE Confidence: 0.97096187

00:10:25.280 --> 00:10:26.480 a number of other foundation
NOTE Confidence: 0.97096187

00:10:26.480 --> 00:10:27.360 models and a number of
NOTE Confidence: 0.97096187

00:10:27.360 --> 00:10:27.860 other,
NOTE Confidence: 0.9346503

00:10:28.960 --> 00:10:30.320 standardized models that were commonly
NOTE Confidence: 0.9346503

00:10:30.320 --> 00:10:31.679 used including the ResNet fifty
NOTE Confidence: 0.9346503

00:10:31.679 --> 00:10:32.975 that's just trained on ImageNet,
NOTE Confidence: 0.9346503

00:10:33.195 --> 00:10:34.495 which was the most commonly

NOTE Confidence: 0.9913563
00:10:35.115 --> 00:10:36.875 used model until until very
NOTE Confidence: 0.9913563
00:10:36.875 --> 00:10:37.375 recently.
NOTE Confidence: 0.9670286
00:10:38.154 --> 00:10:39.515 We apply this to thirty
NOTE Confidence: 0.9670286
00:10:39.515 --> 00:10:41.135 three downstream tasks including
NOTE Confidence: 0.7706684
00:10:41.515 --> 00:10:42.015 for
NOTE Confidence: 0.93916357
00:10:42.635 --> 00:10:45.295 ROI level classification, segmentation, retrieval,
NOTE Confidence: 0.9731695
00:10:46.220 --> 00:10:47.019 at the level of the
NOTE Confidence: 0.9731695
00:10:47.019 --> 00:10:48.059 whole slide and at the
NOTE Confidence: 0.9731695
00:10:48.059 --> 00:10:49.300 level of regions of interest.
NOTE Confidence: 0.9911789
00:10:49.740 --> 00:10:51.259 This radar plot is a
NOTE Confidence: 0.9911789
00:10:51.259 --> 00:10:52.639 really bad way to present
NOTE Confidence: 0.9911789
00:10:52.699 --> 00:10:54.620 all of these datasets together,
NOTE Confidence: 0.9911789
00:10:54.620 --> 00:10:55.740 but it does let you
NOTE Confidence: 0.9911789
00:10:55.740 --> 00:10:57.519 see all the different tasks,
NOTE Confidence: 0.96112293
00:10:58.459 --> 00:10:59.579 that we apply this to
NOTE Confidence: 0.96112293

00:10:59.579 --> 00:11:00.925 in one go. The bar
NOTE Confidence: 0.96112293

00:11:00.925 --> 00:11:01.805 plots are a much better
NOTE Confidence: 0.96112293

00:11:01.805 --> 00:11:03.485 way to statistically observe the
NOTE Confidence: 0.96112293

00:11:03.485 --> 00:11:04.625 improvement in performance,
NOTE Confidence: 0.98726916

00:11:04.925 --> 00:11:06.045 and we showed that there
NOTE Confidence: 0.98726916

00:11:06.045 --> 00:11:07.085 was about a six to
NOTE Confidence: 0.98726916

00:11:07.085 --> 00:11:08.225 eight percent improvement,
NOTE Confidence: 0.9890147

00:11:08.925 --> 00:11:10.225 in performance over,
NOTE Confidence: 0.9950775

00:11:10.765 --> 00:11:11.725 some of the other models
NOTE Confidence: 0.9950775

00:11:11.725 --> 00:11:13.505 and a substantial improvement over,
NOTE Confidence: 0.9718335

00:11:14.205 --> 00:11:15.720 or a ResNet fifty based
NOTE Confidence: 0.9718335

00:11:15.959 --> 00:11:17.720 based model. And perhaps the
NOTE Confidence: 0.9718335

00:11:17.720 --> 00:11:19.319 more clinically useful aspect of
NOTE Confidence: 0.9718335

00:11:19.319 --> 00:11:20.360 this is around few shot
NOTE Confidence: 0.9718335

00:11:20.360 --> 00:11:21.240 learning where you can use
NOTE Confidence: 0.9718335

00:11:21.240 --> 00:11:22.279 very few data points to

NOTE Confidence: 0.9718335
00:11:22.279 --> 00:11:23.819 train models that,
NOTE Confidence: 0.99726266
00:11:24.519 --> 00:11:26.540 that can have clinical downstream
NOTE Confidence: 0.99726266
00:11:26.600 --> 00:11:27.420 clinical applicability.
NOTE Confidence: 0.95733464
00:11:28.395 --> 00:11:29.675 In parallel, we also worked
NOTE Confidence: 0.95733464
00:11:29.675 --> 00:11:31.515 on a image text model.
NOTE Confidence: 0.95733464
00:11:31.515 --> 00:11:32.875 So trying to see if
NOTE Confidence: 0.95733464
00:11:32.875 --> 00:11:34.235 we can contrast with text
NOTE Confidence: 0.95733464
00:11:34.235 --> 00:11:35.995 to improve improve image based
NOTE Confidence: 0.95733464
00:11:35.995 --> 00:11:36.895 feature representations.
NOTE Confidence: 0.9764085
00:11:37.355 --> 00:11:38.315 And we showed that you
NOTE Confidence: 0.9764085
00:11:38.315 --> 00:11:39.835 could use this model in
NOTE Confidence: 0.9764085
00:11:39.835 --> 00:11:40.875 a variety of different ways.
NOTE Confidence: 0.9764085
00:11:40.875 --> 00:11:41.935 This is just a distribution
NOTE Confidence: 0.9764085
00:11:41.995 --> 00:11:43.230 of all the data that
NOTE Confidence: 0.9764085
00:11:43.230 --> 00:11:44.429 was used for this. Most
NOTE Confidence: 0.9764085

00:11:44.429 --> 00:11:45.230 of the data came from
NOTE Confidence: 0.9764085

00:11:45.230 --> 00:11:46.750 the PubMed open access database
NOTE Confidence: 0.9764085

00:11:46.750 --> 00:11:48.350 where we use human pathology
NOTE Confidence: 0.9764085

00:11:48.350 --> 00:11:50.110 images and corresponding captions often
NOTE Confidence: 0.9764085

00:11:50.110 --> 00:11:51.550 linking with what
NOTE Confidence: 0.9950684

00:11:52.030 --> 00:11:53.390 where those images are referred
NOTE Confidence: 0.9950684

00:11:53.390 --> 00:11:54.510 to within the within the
NOTE Confidence: 0.9950684

00:11:54.510 --> 00:11:56.190 article to expand on those
NOTE Confidence: 0.9950684

00:11:56.190 --> 00:11:56.929 text captions.
NOTE Confidence: 0.9882303

00:11:57.355 --> 00:11:58.554 But the richness of these
NOTE Confidence: 0.9882303

00:11:58.554 --> 00:11:59.675 features were assessed in a
NOTE Confidence: 0.9882303

00:11:59.675 --> 00:12:01.054 number of different ways, including
NOTE Confidence: 0.9628839

00:12:01.434 --> 00:12:02.955 with zero shot classification where
NOTE Confidence: 0.9628839

00:12:02.955 --> 00:12:03.835 the goal is not to
NOTE Confidence: 0.9628839

00:12:03.835 --> 00:12:05.675 have clinically useful models, but
NOTE Confidence: 0.9628839

00:12:05.675 --> 00:12:07.275 to see raw features, how

NOTE Confidence: 0.9628839
00:12:07.355 --> 00:12:08.395 just how rich the raw
NOTE Confidence: 0.9628839
00:12:08.395 --> 00:12:09.910 features are. And then they
NOTE Confidence: 0.9628839
00:12:09.910 --> 00:12:10.710 could be used in a
NOTE Confidence: 0.9628839
00:12:10.710 --> 00:12:12.309 supervised setting for within a
NOTE Confidence: 0.9628839
00:12:12.309 --> 00:12:13.910 few shot learning or other
NOTE Confidence: 0.9628839
00:12:13.910 --> 00:12:14.870 kinds of tasks. In this
NOTE Confidence: 0.9628839
00:12:14.870 --> 00:12:16.150 example, we're showing non small
NOTE Confidence: 0.9628839
00:12:16.150 --> 00:12:17.290 cell lung cancer subtyping
NOTE Confidence: 0.9543397
00:12:17.750 --> 00:12:18.870 by using very, very few
NOTE Confidence: 0.9543397
00:12:18.870 --> 00:12:20.250 samples, how well the model
NOTE Confidence: 0.9543397
00:12:20.309 --> 00:12:21.670 can perform with just using
NOTE Confidence: 0.9543397
00:12:21.670 --> 00:12:23.030 four samples or eight samples
NOTE Confidence: 0.9543397
00:12:23.030 --> 00:12:23.445 for
NOTE Confidence: 0.9719685
00:12:23.845 --> 00:12:25.045 for for training and whether
NOTE Confidence: 0.9719685
00:12:25.045 --> 00:12:25.925 we can get to clinically
NOTE Confidence: 0.9719685

00:12:25.925 --> 00:12:26.665 useful performance
NOTE Confidence: 0.96807736

00:12:26.965 --> 00:12:28.245 by using fewer samples just
NOTE Confidence: 0.96807736

00:12:28.245 --> 00:12:29.125 by which of the fact
NOTE Confidence: 0.96807736

00:12:29.125 --> 00:12:30.325 that the features are much
NOTE Confidence: 0.96807736

00:12:30.325 --> 00:12:31.545 more rich,
NOTE Confidence: 0.9333391

00:12:32.165 --> 00:12:33.285 at this point. So we
NOTE Confidence: 0.9333391

00:12:33.285 --> 00:12:34.505 made both of these models,
NOTE Confidence: 0.9333391

00:12:34.725 --> 00:12:37.179 publicly available, and we've had,
NOTE Confidence: 0.9333391

00:12:37.179 --> 00:12:38.220 you know, or or four
NOTE Confidence: 0.9333391

00:12:38.220 --> 00:12:40.220 hundred thousand downloads over or
NOTE Confidence: 0.9333391

00:12:40.220 --> 00:12:41.339 hugging face. And the models
NOTE Confidence: 0.9333391

00:12:41.339 --> 00:12:42.380 have been used around the
NOTE Confidence: 0.9333391

00:12:42.380 --> 00:12:43.259 around the world for a
NOTE Confidence: 0.9333391

00:12:43.259 --> 00:12:45.100 variety of different tasks. And
NOTE Confidence: 0.9333391

00:12:45.100 --> 00:12:46.300 they continue to be used
NOTE Confidence: 0.9333391

00:12:46.300 --> 00:12:47.820 for a number of different

NOTE Confidence: 0.9333391

00:12:47.820 --> 00:12:48.660 reasons. There there are there

NOTE Confidence: 0.9333391

00:12:48.660 --> 00:12:49.679 are there are a few

NOTE Confidence: 0.94311774

00:12:50.875 --> 00:12:52.554 independent analysis. This analysis is

NOTE Confidence: 0.94311774

00:12:52.554 --> 00:12:53.615 from Jacob Cather's

NOTE Confidence: 0.91613257

00:12:54.074 --> 00:12:55.755 group, in Dresden where they

NOTE Confidence: 0.91613257

00:12:55.755 --> 00:12:57.115 showed that both Unni and

NOTE Confidence: 0.91613257

00:12:57.115 --> 00:12:58.875 Conch performed quite well on

NOTE Confidence: 0.91613257

00:12:58.875 --> 00:12:59.855 a number of these,

NOTE Confidence: 0.98008364

00:13:00.875 --> 00:13:02.554 tasks, both using publicly available

NOTE Confidence: 0.98008364

00:13:02.554 --> 00:13:03.355 data as well as some

NOTE Confidence: 0.98008364

00:13:03.355 --> 00:13:04.660 of the internal data that

NOTE Confidence: 0.98008364

00:13:04.660 --> 00:13:05.540 they that they had. And

NOTE Confidence: 0.98008364

00:13:05.540 --> 00:13:06.840 we believe that the diversity

NOTE Confidence: 0.96342415

00:13:07.300 --> 00:13:08.340 of data that was used

NOTE Confidence: 0.96342415

00:13:08.340 --> 00:13:10.200 for training, which included infectious

NOTE Confidence: 0.96342415

00:13:10.260 --> 00:13:12.120 inflammatory and neoplastic cases,
NOTE Confidence: 0.9885416

00:13:12.660 --> 00:13:14.260 is, is what's leading to
NOTE Confidence: 0.9885416

00:13:14.260 --> 00:13:15.080 the improved,
NOTE Confidence: 0.995494

00:13:15.540 --> 00:13:17.755 improved performance. There's some additional
NOTE Confidence: 0.995494

00:13:17.755 --> 00:13:19.355 analysis. This this analysis was
NOTE Confidence: 0.995494

00:13:19.355 --> 00:13:20.334 from Mount Sinai,
NOTE Confidence: 0.9907159

00:13:21.115 --> 00:13:22.255 where they showed that
NOTE Confidence: 0.9891027

00:13:23.115 --> 00:13:24.075 the the model did quite
NOTE Confidence: 0.9891027

00:13:24.075 --> 00:13:24.815 well for
NOTE Confidence: 0.94464296

00:13:25.515 --> 00:13:26.635 just a number of different
NOTE Confidence: 0.94464296

00:13:26.635 --> 00:13:27.675 tasks. So I'll refer you
NOTE Confidence: 0.94464296

00:13:27.675 --> 00:13:29.515 to these these studies for
NOTE Confidence: 0.94464296

00:13:29.515 --> 00:13:30.875 for a more in-depth analysis.
NOTE Confidence: 0.94464296

00:13:30.875 --> 00:13:32.154 So some analysis that I
NOTE Confidence: 0.94464296

00:13:32.154 --> 00:13:33.900 like to show is around
NOTE Confidence: 0.94464296

00:13:33.900 --> 00:13:34.720 how fast

NOTE Confidence: 0.9855534
00:13:35.580 --> 00:13:37.260 the the feature extraction can
NOTE Confidence: 0.9855534
00:13:37.260 --> 00:13:38.220 be from some of these
NOTE Confidence: 0.9855534
00:13:38.220 --> 00:13:38.720 models,
NOTE Confidence: 0.88796717
00:13:40.059 --> 00:13:41.500 and what the storage cost
NOTE Confidence: 0.88796717
00:13:41.500 --> 00:13:42.540 would be because that's how
NOTE Confidence: 0.88796717
00:13:42.540 --> 00:13:43.740 you we we can practically
NOTE Confidence: 0.88796717
00:13:43.740 --> 00:13:45.260 use this. So there there
NOTE Confidence: 0.88796717
00:13:45.260 --> 00:13:46.160 are there are two
NOTE Confidence: 0.999223
00:13:47.745 --> 00:13:48.245 possible
NOTE Confidence: 0.9854546
00:13:48.785 --> 00:13:49.745 ways we can use these
NOTE Confidence: 0.9854546
00:13:49.745 --> 00:13:50.625 features. We can use these
NOTE Confidence: 0.9854546
00:13:50.625 --> 00:13:52.705 features to train new models.
NOTE Confidence: 0.9854546
00:13:52.705 --> 00:13:54.385 So every slide that scanned,
NOTE Confidence: 0.9854546
00:13:54.385 --> 00:13:55.345 if you extract features and
NOTE Confidence: 0.9854546
00:13:55.345 --> 00:13:56.005 keep them,
NOTE Confidence: 0.96475005

00:13:56.465 --> 00:13:57.105 it can be used to
NOTE Confidence: 0.96475005

00:13:57.105 --> 00:13:58.225 train new models, but they
NOTE Confidence: 0.96475005

00:13:58.225 --> 00:13:59.105 can also be used for
NOTE Confidence: 0.96475005

00:13:59.105 --> 00:14:01.265 model inference. So if you
NOTE Confidence: 0.96475005

00:14:01.265 --> 00:14:02.885 extract these rich feature representations
NOTE Confidence: 0.9523196

00:14:03.590 --> 00:14:04.970 from the whole slide image
NOTE Confidence: 0.9523196

00:14:05.270 --> 00:14:06.390 and then store them, you
NOTE Confidence: 0.9523196

00:14:06.390 --> 00:14:07.510 can use them both for
NOTE Confidence: 0.9523196

00:14:07.510 --> 00:14:08.710 model development as well as
NOTE Confidence: 0.9523196

00:14:08.710 --> 00:14:10.490 for model model inference.
NOTE Confidence: 0.9624575

00:14:10.790 --> 00:14:11.830 And both UNI and Conch
NOTE Confidence: 0.9624575

00:14:11.830 --> 00:14:12.710 do quite well in that
NOTE Confidence: 0.9624575

00:14:12.710 --> 00:14:13.750 in that regard. We already
NOTE Confidence: 0.9624575

00:14:13.750 --> 00:14:14.250 have
NOTE Confidence: 0.98139745

00:14:15.270 --> 00:14:16.309 a version of UNI two
NOTE Confidence: 0.98139745

00:14:16.309 --> 00:14:17.605 and Conch two that we

NOTE Confidence: 0.98139745
00:14:17.605 --> 00:14:18.725 intend to make public in
NOTE Confidence: 0.98139745
00:14:18.725 --> 00:14:19.945 the in the coming days.
NOTE Confidence: 0.9824069
00:14:21.525 --> 00:14:22.885 In parallel, we have been
NOTE Confidence: 0.9824069
00:14:22.885 --> 00:14:24.325 working on a slide level
NOTE Confidence: 0.9824069
00:14:24.325 --> 00:14:25.685 foundation model where the goal
NOTE Confidence: 0.9824069
00:14:25.685 --> 00:14:27.865 is to extract a single,
NOTE Confidence: 0.9876516
00:14:29.125 --> 00:14:29.525 feature
NOTE Confidence: 0.9770905
00:14:30.165 --> 00:14:31.800 a single feature vector corresponding
NOTE Confidence: 0.9770905
00:14:31.800 --> 00:14:33.000 a whole whole slide image.
NOTE Confidence: 0.9770905
00:14:33.000 --> 00:14:33.640 And it could be used
NOTE Confidence: 0.9770905
00:14:33.640 --> 00:14:34.779 for a number of different
NOTE Confidence: 0.92447144
00:14:37.000 --> 00:14:39.000 tasks including for retrieval at
NOTE Confidence: 0.92447144
00:14:39.000 --> 00:14:39.800 the level of the whole
NOTE Confidence: 0.92447144
00:14:39.800 --> 00:14:40.300 slide,
NOTE Confidence: 0.9979925
00:14:41.000 --> 00:14:41.500 and
NOTE Confidence: 0.9494996

00:14:42.040 --> 00:14:44.595 for very simple classification problems.

NOTE Confidence: 0.9494996

00:14:44.595 --> 00:14:45.735 It could it could turn,

NOTE Confidence: 0.9991303

00:14:46.274 --> 00:14:47.575 a lot of these complicated

NOTE Confidence: 0.9991303

00:14:47.795 --> 00:14:48.295 classification

NOTE Confidence: 0.97338134

00:14:48.834 --> 00:14:50.375 pipelines into a very simplistic,

NOTE Confidence: 0.9709056

00:14:51.315 --> 00:14:52.755 classification pipeline if you if

NOTE Confidence: 0.9709056

00:14:52.755 --> 00:14:54.195 you have the single feature

NOTE Confidence: 0.9709056

00:14:54.195 --> 00:14:55.394 vector corresponding the whole slide

NOTE Confidence: 0.9709056

00:14:55.394 --> 00:14:56.670 image. So we're using the

NOTE Confidence: 0.9709056

00:14:56.829 --> 00:14:57.870 whole slide images for this,

NOTE Confidence: 0.9709056

00:14:57.870 --> 00:14:59.070 but we're also contrasting with

NOTE Confidence: 0.9709056

00:14:59.070 --> 00:14:59.570 text.

NOTE Confidence: 0.9933629

00:15:00.110 --> 00:15:01.310 In this case, we're contrasting

NOTE Confidence: 0.9933629

00:15:01.310 --> 00:15:02.670 with text that came from

NOTE Confidence: 0.9933629

00:15:02.670 --> 00:15:03.870 a generative AI model that

NOTE Confidence: 0.9933629

00:15:03.870 --> 00:15:04.610 we developed

NOTE Confidence: 0.98085344
00:15:04.990 --> 00:15:07.410 and, the pathology pathology report.
NOTE Confidence: 0.95237267
00:15:08.670 --> 00:15:09.790 And we we we've shown
NOTE Confidence: 0.95237267
00:15:09.790 --> 00:15:11.325 that this form of contrasting
NOTE Confidence: 0.95237267
00:15:11.465 --> 00:15:13.225 leads to substantial improvements or
NOTE Confidence: 0.95237267
00:15:13.225 --> 00:15:14.265 some of the some of
NOTE Confidence: 0.95237267
00:15:14.265 --> 00:15:15.965 the other models for morphologic
NOTE Confidence: 0.95237267
00:15:16.105 --> 00:15:17.485 subtyping, IHC quantification,
NOTE Confidence: 0.9511161
00:15:18.185 --> 00:15:19.945 biomarker prediction of all of
NOTE Confidence: 0.9511161
00:15:19.945 --> 00:15:21.565 all sorts, and more importantly
NOTE Confidence: 0.9511161
00:15:21.625 --> 00:15:23.160 for few shot classification, which
NOTE Confidence: 0.9511161
00:15:23.160 --> 00:15:24.520 is where the real clinical
NOTE Confidence: 0.9511161
00:15:24.520 --> 00:15:25.880 utility lies, where you can
NOTE Confidence: 0.9511161
00:15:25.880 --> 00:15:27.900 use very, very few examples
NOTE Confidence: 0.9511161
00:15:28.200 --> 00:15:28.700 for
NOTE Confidence: 0.86363816
00:15:29.240 --> 00:15:30.760 or or very, very few
NOTE Confidence: 0.86363816

00:15:30.760 --> 00:15:32.600 images for training some of
NOTE Confidence: 0.86363816

00:15:32.600 --> 00:15:33.100 these,
NOTE Confidence: 0.9986011

00:15:33.720 --> 00:15:34.780 some of these models.
NOTE Confidence: 0.9882047

00:15:35.585 --> 00:15:37.025 And we've tested this on
NOTE Confidence: 0.9882047

00:15:37.025 --> 00:15:38.145 a number of different, right,
NOTE Confidence: 0.9882047

00:15:38.145 --> 00:15:38.645 difficult
NOTE Confidence: 0.86777836

00:15:39.825 --> 00:15:41.905 cases including for the for
NOTE Confidence: 0.86777836

00:15:41.905 --> 00:15:42.405 classifying
NOTE Confidence: 0.90801156

00:15:43.105 --> 00:15:44.545 all the brain tumors in
NOTE Confidence: 0.90801156

00:15:44.545 --> 00:15:45.605 the eBrains dataset,
NOTE Confidence: 0.99760914

00:15:46.145 --> 00:15:47.185 as well as for treatment
NOTE Confidence: 0.99760914

00:15:47.185 --> 00:15:48.085 response prediction
NOTE Confidence: 0.95981497

00:15:49.290 --> 00:15:50.649 task. We'll we'll be hopefully
NOTE Confidence: 0.95981497

00:15:50.649 --> 00:15:52.010 putting this preprint out very
NOTE Confidence: 0.95981497

00:15:52.010 --> 00:15:53.709 soon. Because it was contrasted
NOTE Confidence: 0.95981497

00:15:53.769 --> 00:15:55.450 with text, we could also

NOTE Confidence: 0.95981497
00:15:55.450 --> 00:15:56.350 look at,
NOTE Confidence: 0.96951115
00:15:57.450 --> 00:15:58.810 how well it does in
NOTE Confidence: 0.96951115
00:15:58.810 --> 00:16:00.350 terms of zero shot classification,
NOTE Confidence: 0.9932791
00:16:01.625 --> 00:16:03.485 and for generating the report,
NOTE Confidence: 0.93382466
00:16:03.945 --> 00:16:05.385 directly from the from the
NOTE Confidence: 0.93382466
00:16:05.385 --> 00:16:07.165 pathology image or a multitude
NOTE Confidence: 0.93382466
00:16:07.225 --> 00:16:09.465 of of pathology images. And
NOTE Confidence: 0.93382466
00:16:09.465 --> 00:16:10.585 we tested this on, you
NOTE Confidence: 0.93382466
00:16:10.585 --> 00:16:12.205 know, basically the entire TCGA
NOTE Confidence: 0.93382466
00:16:12.425 --> 00:16:13.465 on a on a internal
NOTE Confidence: 0.93382466
00:16:13.465 --> 00:16:14.905 dataset that we call OT
NOTE Confidence: 0.93382466
00:16:14.905 --> 00:16:15.900 one zero eight, which is,
NOTE Confidence: 0.93985736
00:16:16.860 --> 00:16:17.660 a group of hundred and
NOTE Confidence: 0.93985736
00:16:17.660 --> 00:16:18.720 eight difficult diagnoses
NOTE Confidence: 0.9045253
00:16:19.820 --> 00:16:20.780 as well as on on
NOTE Confidence: 0.9045253

00:16:20.780 --> 00:16:22.320 ebrains and other other datasets.
NOTE Confidence: 0.9661027

00:16:23.340 --> 00:16:24.860 We plan to release this
NOTE Confidence: 0.9661027

00:16:24.860 --> 00:16:26.640 model, in the coming weeks.
NOTE Confidence: 0.9661027

00:16:26.940 --> 00:16:28.300 So people who are interested
NOTE Confidence: 0.9661027

00:16:28.300 --> 00:16:29.360 in this, stay tuned.
NOTE Confidence: 0.9369701

00:16:31.035 --> 00:16:31.435 We
NOTE Confidence: 0.96170413

00:16:31.995 --> 00:16:33.115 so so the story so
NOTE Confidence: 0.96170413

00:16:33.115 --> 00:16:33.995 far is that we've shown
NOTE Confidence: 0.96170413

00:16:33.995 --> 00:16:34.954 that you can extract which
NOTE Confidence: 0.96170413

00:16:34.954 --> 00:16:37.134 feature representations from pathology images,
NOTE Confidence: 0.968206

00:16:37.675 --> 00:16:39.134 and there were a multitude
NOTE Confidence: 0.968206

00:16:39.194 --> 00:16:40.634 of self supervised models around
NOTE Confidence: 0.968206

00:16:40.634 --> 00:16:42.074 that. And we can contrast
NOTE Confidence: 0.968206

00:16:42.074 --> 00:16:43.694 with text and improve the
NOTE Confidence: 0.968206

00:16:43.834 --> 00:16:45.740 the feature representation. But there
NOTE Confidence: 0.968206

00:16:45.740 --> 00:16:47.500 are other, modalities that we

NOTE Confidence: 0.968206
00:16:47.500 --> 00:16:49.019 have about around these cases
NOTE Confidence: 0.968206
00:16:49.019 --> 00:16:50.060 that we can contrast with
NOTE Confidence: 0.968206
00:16:50.060 --> 00:16:52.000 to further improve feature presentation.
NOTE Confidence: 0.968206
00:16:52.140 --> 00:16:53.820 And in this particular case,
NOTE Confidence: 0.968206
00:16:53.820 --> 00:16:55.420 we're contrasting with IHCs. This
NOTE Confidence: 0.968206
00:16:55.420 --> 00:16:56.079 is a,
NOTE Confidence: 0.9904846
00:16:56.779 --> 00:16:58.060 article that was just presented
NOTE Confidence: 0.9904846
00:16:58.060 --> 00:16:58.880 at ECCV,
NOTE Confidence: 0.93633956
00:17:00.185 --> 00:17:02.365 where we contrast pathology images
NOTE Confidence: 0.93633956
00:17:02.425 --> 00:17:04.284 with, with with the IHCs,
NOTE Confidence: 0.9845291
00:17:05.865 --> 00:17:06.825 and show that you can
NOTE Confidence: 0.9845291
00:17:06.825 --> 00:17:08.424 improve each representation and improve
NOTE Confidence: 0.9845291
00:17:08.424 --> 00:17:10.585 IHC quantification without requiring any
NOTE Confidence: 0.9845291
00:17:10.585 --> 00:17:11.645 pixel level annotation,
NOTE Confidence: 0.95177716
00:17:12.585 --> 00:17:13.880 as as well as other
NOTE Confidence: 0.95177716

00:17:13.880 --> 00:17:15.320 cases like survival and so
NOTE Confidence: 0.95177716

00:17:15.320 --> 00:17:16.360 forth. So if you can
NOTE Confidence: 0.95177716

00:17:16.360 --> 00:17:17.400 contrast with text and you
NOTE Confidence: 0.95177716

00:17:17.400 --> 00:17:19.160 can contrast with the administered
NOTE Confidence: 0.95177716

00:17:19.160 --> 00:17:20.360 chemistry, you can also contrast
NOTE Confidence: 0.95177716

00:17:20.360 --> 00:17:21.020 with transcriptomics.
NOTE Confidence: 0.9973224

00:17:21.560 --> 00:17:22.520 And that's what we showed
NOTE Confidence: 0.9973224

00:17:22.520 --> 00:17:23.340 in this particular
NOTE Confidence: 0.9351568

00:17:23.720 --> 00:17:24.760 study. It was published at
NOTE Confidence: 0.9351568

00:17:24.760 --> 00:17:25.260 CVPR,
NOTE Confidence: 0.9841358

00:17:26.280 --> 00:17:27.420 earlier this year,
NOTE Confidence: 0.9528868

00:17:27.880 --> 00:17:29.025 where we showed that that
NOTE Confidence: 0.9528868

00:17:29.025 --> 00:17:29.984 this was limited to the
NOTE Confidence: 0.9528868

00:17:29.984 --> 00:17:31.425 TCGA where we show contrasting
NOTE Confidence: 0.9528868

00:17:31.425 --> 00:17:32.945 H and E images with
NOTE Confidence: 0.9528868

00:17:32.945 --> 00:17:35.345 the corresponding bulk transcriptomic profile

NOTE Confidence: 0.9528868
00:17:35.345 --> 00:17:37.205 can improve few shot classification
NOTE Confidence: 0.99732417
00:17:37.984 --> 00:17:39.125 on lung cancers,
NOTE Confidence: 0.9519435
00:17:39.744 --> 00:17:41.345 breast cancers, as well as
NOTE Confidence: 0.9519435
00:17:41.345 --> 00:17:42.565 for, toxicology.
NOTE Confidence: 0.99182945
00:17:43.850 --> 00:17:45.290 More recently, we've expanded this
NOTE Confidence: 0.99182945
00:17:45.290 --> 00:17:46.170 to use all of the
NOTE Confidence: 0.99182945
00:17:46.170 --> 00:17:47.770 molecular data that was available
NOTE Confidence: 0.99182945
00:17:47.770 --> 00:17:49.050 at Brigham and MGH. So
NOTE Confidence: 0.99182945
00:17:49.050 --> 00:17:49.550 combining
NOTE Confidence: 0.9924276
00:17:50.170 --> 00:17:50.830 all the,
NOTE Confidence: 0.9923992
00:17:52.730 --> 00:17:53.630 all the transcriptomic
NOTE Confidence: 0.9464023
00:17:54.170 --> 00:17:55.130 data that we have at
NOTE Confidence: 0.9464023
00:17:55.130 --> 00:17:56.655 at MGH and all the
NOTE Confidence: 0.9464023
00:17:56.734 --> 00:17:58.015 NGS data that we have
NOTE Confidence: 0.9464023
00:17:58.015 --> 00:17:59.055 at the, at the Brigham
NOTE Confidence: 0.9464023

00:17:59.055 --> 00:18:00.895 and Dana Farber and contrasting
NOTE Confidence: 0.9464023

00:18:00.895 --> 00:18:02.095 with that to see if
NOTE Confidence: 0.9464023

00:18:02.095 --> 00:18:03.555 we can improve feature representation.
NOTE Confidence: 0.9464023

00:18:03.695 --> 00:18:04.655 Then we apply this to
NOTE Confidence: 0.9464023

00:18:04.655 --> 00:18:06.255 a number of different downstream
NOTE Confidence: 0.9464023

00:18:06.255 --> 00:18:08.195 tasks from mutation prediction, from
NOTE Confidence: 0.9464023

00:18:08.255 --> 00:18:09.935 from HNE images to molecular
NOTE Confidence: 0.9464023

00:18:09.935 --> 00:18:11.670 subtyping, and more importantly for
NOTE Confidence: 0.9464023

00:18:11.830 --> 00:18:13.190 treatment response predictions. If you
NOTE Confidence: 0.9464023

00:18:13.190 --> 00:18:14.150 look at some of the
NOTE Confidence: 0.9464023

00:18:14.150 --> 00:18:15.350 results around here for treatment
NOTE Confidence: 0.9464023

00:18:15.350 --> 00:18:16.170 response prediction,
NOTE Confidence: 0.9652377

00:18:17.109 --> 00:18:18.710 there's a substantial improvement over
NOTE Confidence: 0.9652377

00:18:18.710 --> 00:18:20.550 image based image based models
NOTE Confidence: 0.9652377

00:18:20.550 --> 00:18:22.310 and image text image text
NOTE Confidence: 0.9652377

00:18:22.310 --> 00:18:25.109 models because transcriptomics represents represents

NOTE Confidence: 0.9652377

00:18:25.109 --> 00:18:26.390 a form of contrasting that's

NOTE Confidence: 0.9652377

00:18:26.390 --> 00:18:27.130 much richer

NOTE Confidence: 0.97854805

00:18:27.575 --> 00:18:28.775 in comparison to text. And

NOTE Confidence: 0.97854805

00:18:28.775 --> 00:18:30.555 we hope that combining transcriptomics

NOTE Confidence: 0.97854805

00:18:30.695 --> 00:18:32.455 and text and images leads

NOTE Confidence: 0.97854805

00:18:32.455 --> 00:18:33.835 to an even even,

NOTE Confidence: 0.9543646

00:18:34.615 --> 00:18:35.755 richer feature representation.

NOTE Confidence: 0.9400048

00:18:37.975 --> 00:18:38.375 So,

NOTE Confidence: 0.9849352

00:18:39.015 --> 00:18:41.030 we'll make this preprint publicly

NOTE Confidence: 0.9849352

00:18:41.030 --> 00:18:42.630 available very soon. So the

NOTE Confidence: 0.9849352

00:18:42.630 --> 00:18:43.990 story so far is that

NOTE Confidence: 0.9849352

00:18:43.990 --> 00:18:44.869 we have shown that we

NOTE Confidence: 0.9849352

00:18:44.869 --> 00:18:45.609 can contrast,

NOTE Confidence: 0.90650725

00:18:46.390 --> 00:18:47.510 so so we we can

NOTE Confidence: 0.90650725

00:18:47.510 --> 00:18:49.109 build large self supervised models
NOTE Confidence: 0.90650725

00:18:49.109 --> 00:18:49.930 based on histology.
NOTE Confidence: 0.9993352

00:18:50.790 --> 00:18:51.910 We can contrast them with
NOTE Confidence: 0.9993352

00:18:51.910 --> 00:18:52.410 text
NOTE Confidence: 0.9885059

00:18:52.744 --> 00:18:54.505 and get retrofit representations. We
NOTE Confidence: 0.9885059

00:18:54.505 --> 00:18:55.725 can contrast with IHC
NOTE Confidence: 0.9466423

00:18:56.105 --> 00:18:57.885 to get rich rich representations,
NOTE Confidence: 0.9466423

00:18:58.025 --> 00:18:59.385 and we can contrast with
NOTE Confidence: 0.9466423

00:18:59.385 --> 00:19:00.285 with with transcriptomics,
NOTE Confidence: 0.980145

00:19:02.585 --> 00:19:04.365 to continue to improve,
NOTE Confidence: 0.9893454

00:19:04.744 --> 00:19:06.445 the representation of the image
NOTE Confidence: 0.9893454

00:19:06.700 --> 00:19:07.660 in terms of its its
NOTE Confidence: 0.9893454

00:19:07.660 --> 00:19:09.020 its features. But what can
NOTE Confidence: 0.9893454

00:19:09.020 --> 00:19:09.980 we do with these features?
NOTE Confidence: 0.9893454

00:19:09.980 --> 00:19:10.460 There are a number of
NOTE Confidence: 0.9893454

00:19:10.460 --> 00:19:11.340 different things that can be

NOTE Confidence: 0.9893454
00:19:11.340 --> 00:19:12.380 done. Of course, you can
NOTE Confidence: 0.9893454
00:19:12.380 --> 00:19:14.240 use it to improve the
NOTE Confidence: 0.70085245
00:19:16.060 --> 00:19:17.740 Still sharing on Zoom. Someone's
NOTE Confidence: 0.70085245
00:19:17.740 --> 00:19:18.700 saying that we we can
NOTE Confidence: 0.70085245
00:19:18.700 --> 00:19:19.680 all see the part.
NOTE Confidence: 0.93652624
00:19:20.585 --> 00:19:22.345 So we're having technical issues
NOTE Confidence: 0.93652624
00:19:22.345 --> 00:19:23.465 with Zoom. Why don't you
NOTE Confidence: 0.93652624
00:19:23.465 --> 00:19:24.925 continue on? Okay.
NOTE Confidence: 0.8193456
00:19:25.945 --> 00:19:26.445 So,
NOTE Confidence: 0.8484535
00:19:28.905 --> 00:19:29.405 so,
NOTE Confidence: 0.9852119
00:19:30.984 --> 00:19:32.025 we've shown that you can
NOTE Confidence: 0.9852119
00:19:32.025 --> 00:19:33.545 get to richer richer feature
NOTE Confidence: 0.9852119
00:19:33.545 --> 00:19:34.905 representations using all of these
NOTE Confidence: 0.9852119
00:19:34.905 --> 00:19:36.640 different contrastive methods. But what
NOTE Confidence: 0.9852119
00:19:36.640 --> 00:19:37.440 can you do with these
NOTE Confidence: 0.9852119

00:19:37.440 --> 00:19:39.200 richer feature representations? You can
NOTE Confidence: 0.9852119

00:19:39.200 --> 00:19:40.020 build better
NOTE Confidence: 0.9598653

00:19:40.559 --> 00:19:43.140 supervised models, targeted supervised models.
NOTE Confidence: 0.9598653

00:19:43.359 --> 00:19:44.640 But in parallel, there's been
NOTE Confidence: 0.9598653

00:19:44.640 --> 00:19:46.080 all this development in multimodal
NOTE Confidence: 0.9598653

00:19:46.080 --> 00:19:47.220 large language models
NOTE Confidence: 0.9469771

00:19:47.520 --> 00:19:48.400 where you can have a
NOTE Confidence: 0.9469771

00:19:48.400 --> 00:19:49.600 single model that harbors a
NOTE Confidence: 0.9469771

00:19:49.600 --> 00:19:50.740 lot of lot of knowledge.
NOTE Confidence: 0.56380606

00:19:51.575 --> 00:19:52.075 And
NOTE Confidence: 0.9564338

00:19:52.775 --> 00:19:53.734 that's what sort of led
NOTE Confidence: 0.9564338

00:19:53.734 --> 00:19:55.015 to this study because our
NOTE Confidence: 0.9564338

00:19:55.015 --> 00:19:55.675 our hypothesis
NOTE Confidence: 0.9779562

00:19:56.135 --> 00:19:57.494 was was that OpenAI is
NOTE Confidence: 0.9779562

00:19:57.494 --> 00:19:59.015 trying to build a single
NOTE Confidence: 0.9779562

00:19:59.015 --> 00:20:00.295 model that harbors the world's

NOTE Confidence: 0.9779562
00:20:00.295 --> 00:20:00.795 knowledge.
NOTE Confidence: 0.99770176
00:20:01.175 --> 00:20:01.975 So we should be able
NOTE Confidence: 0.99770176
00:20:01.975 --> 00:20:03.015 to build a single model
NOTE Confidence: 0.99770176
00:20:03.015 --> 00:20:04.135 that harbors all of human
NOTE Confidence: 0.99770176
00:20:04.135 --> 00:20:04.635 pathology
NOTE Confidence: 0.9916864
00:20:05.250 --> 00:20:05.970 knowledge. And what do we
NOTE Confidence: 0.9916864
00:20:05.970 --> 00:20:07.730 need to do to essentially
NOTE Confidence: 0.9916864
00:20:07.730 --> 00:20:08.470 get there?
NOTE Confidence: 0.94396514
00:20:08.930 --> 00:20:11.030 In our in our assessment,
NOTE Confidence: 0.94396514
00:20:11.090 --> 00:20:12.530 it's like a rich self
NOTE Confidence: 0.94396514
00:20:12.530 --> 00:20:13.810 supervised model that can extract
NOTE Confidence: 0.94396514
00:20:13.810 --> 00:20:15.970 rich feature feature representations from
NOTE Confidence: 0.94396514
00:20:15.970 --> 00:20:16.630 these images
NOTE Confidence: 0.9723527
00:20:16.945 --> 00:20:18.225 and image text models that
NOTE Confidence: 0.9723527
00:20:18.225 --> 00:20:20.165 can enhance these feature representations
NOTE Confidence: 0.94181114

00:20:20.465 --> 00:20:21.505 based on the based on
NOTE Confidence: 0.94181114

00:20:21.505 --> 00:20:22.465 the text and a large
NOTE Confidence: 0.94181114

00:20:22.465 --> 00:20:23.445 instruction dataset
NOTE Confidence: 0.99550563

00:20:23.905 --> 00:20:24.725 of questions,
NOTE Confidence: 0.96835965

00:20:25.025 --> 00:20:26.325 images, and responses.
NOTE Confidence: 0.9900747

00:20:26.705 --> 00:20:27.905 And eventually, of course, you
NOTE Confidence: 0.9900747

00:20:27.905 --> 00:20:28.645 need robust
NOTE Confidence: 0.9943254

00:20:28.945 --> 00:20:29.445 evaluation.
NOTE Confidence: 0.99188966

00:20:29.910 --> 00:20:31.270 And our philosophy around this
NOTE Confidence: 0.99188966

00:20:31.270 --> 00:20:32.550 is that because pathology images
NOTE Confidence: 0.99188966

00:20:32.550 --> 00:20:34.550 are hierarchical and harbor information
NOTE Confidence: 0.99188966

00:20:34.550 --> 00:20:35.510 at all of these different
NOTE Confidence: 0.99188966

00:20:35.510 --> 00:20:36.869 scales, but we don't have
NOTE Confidence: 0.99188966

00:20:36.869 --> 00:20:38.650 any text information lying around
NOTE Confidence: 0.9907127

00:20:39.109 --> 00:20:39.609 for
NOTE Confidence: 0.9447303

00:20:39.910 --> 00:20:41.130 every scale of,

NOTE Confidence: 0.8681532

00:20:42.070 --> 00:20:42.869 each and every one of

NOTE Confidence: 0.8681532

00:20:42.869 --> 00:20:43.910 these scales. The the the

NOTE Confidence: 0.8681532

00:20:43.910 --> 00:20:45.270 only information we have is

NOTE Confidence: 0.8681532

00:20:45.270 --> 00:20:45.645 essentially

NOTE Confidence: 0.9323289

00:20:56.765 --> 00:20:58.445 understanding of pathology regions at

NOTE Confidence: 0.9323289

00:20:58.445 --> 00:21:00.145 a cellular level leads to

NOTE Confidence: 0.9909547

00:21:00.750 --> 00:21:02.030 slide level and patient level

NOTE Confidence: 0.9909547

00:21:02.030 --> 00:21:02.530 descriptions.

NOTE Confidence: 0.9988905

00:21:02.990 --> 00:21:04.450 So we needed to get

NOTE Confidence: 0.7434235

00:21:05.310 --> 00:21:05.810 annotations.

NOTE Confidence: 0.94975424

00:21:09.070 --> 00:21:10.290 We need to get annotations,

NOTE Confidence: 0.99951005

00:21:10.750 --> 00:21:11.250 at

NOTE Confidence: 0.9947514

00:21:11.710 --> 00:21:12.210 these

NOTE Confidence: 0.98417693

00:21:13.994 --> 00:21:15.835 specific regions within a pathology

NOTE Confidence: 0.98417693

00:21:15.835 --> 00:21:16.875 image that leads to leads

NOTE Confidence: 0.98417693

00:21:16.875 --> 00:21:17.915 to a diagnosis. So that's
NOTE Confidence: 0.98417693

00:21:17.915 --> 00:21:19.275 what we did to collect
NOTE Confidence: 0.98417693

00:21:19.275 --> 00:21:21.595 this very large instruction dataset.
NOTE Confidence: 0.98417693

00:21:21.595 --> 00:21:23.195 This instruction dataset was based
NOTE Confidence: 0.98417693

00:21:23.195 --> 00:21:23.355 on,
NOTE Confidence: 0.95007914

00:21:25.850 --> 00:21:27.369 some training material that we
NOTE Confidence: 0.95007914

00:21:27.369 --> 00:21:28.410 had available at Brigham and
NOTE Confidence: 0.95007914

00:21:28.410 --> 00:21:29.850 MGH, but also lots of
NOTE Confidence: 0.95007914

00:21:29.850 --> 00:21:31.609 manual data curation and then
NOTE Confidence: 0.95007914

00:21:31.609 --> 00:21:33.210 manual curation of guardrails that
NOTE Confidence: 0.95007914

00:21:33.210 --> 00:21:34.330 we needed to build the
NOTE Confidence: 0.95007914

00:21:34.330 --> 00:21:35.770 build the chatbot. So we
NOTE Confidence: 0.95007914

00:21:35.770 --> 00:21:36.970 got to about nine hundred
NOTE Confidence: 0.95007914

00:21:36.970 --> 00:21:38.330 and ninety nine thousand question
NOTE Confidence: 0.95007914

00:21:38.330 --> 00:21:39.765 answer terms that were used
NOTE Confidence: 0.95007914

00:21:39.765 --> 00:21:41.205 to train the, train the

NOTE Confidence: 0.95007914

00:21:41.205 --> 00:21:41.705 chatbot.

NOTE Confidence: 0.9491217

00:21:42.405 --> 00:21:43.605 And, eventually, we we we

NOTE Confidence: 0.9491217

00:21:43.605 --> 00:21:44.645 had a chatbot that that

NOTE Confidence: 0.9491217

00:21:44.645 --> 00:21:45.685 started to work quite well

NOTE Confidence: 0.9491217

00:21:45.685 --> 00:21:46.965 where we could go into

NOTE Confidence: 0.9491217

00:21:46.965 --> 00:21:48.805 pathology image, ask questions about

NOTE Confidence: 0.9491217

00:21:48.805 --> 00:21:50.725 particular regions within the image,

NOTE Confidence: 0.9491217

00:21:50.725 --> 00:21:51.845 and it would give it

NOTE Confidence: 0.9491217

00:21:51.845 --> 00:21:52.950 would give a response. Like,

NOTE Confidence: 0.9491217

00:21:52.950 --> 00:21:53.830 so so in in in

NOTE Confidence: 0.9491217

00:21:53.830 --> 00:21:55.429 this example, the user is

NOTE Confidence: 0.9491217

00:21:55.429 --> 00:21:56.330 basically asking,

NOTE Confidence: 0.9425882

00:21:56.630 --> 00:21:57.990 you know, what's what what

NOTE Confidence: 0.9425882

00:21:57.990 --> 00:21:58.950 what type of tumor do

NOTE Confidence: 0.9425882

00:21:58.950 --> 00:22:00.070 you see? And as you

NOTE Confidence: 0.9425882

00:22:00.070 --> 00:22:01.510 can see, that limited context
NOTE Confidence: 0.9425882

00:22:01.510 --> 00:22:02.549 was given for this. The
NOTE Confidence: 0.9425882

00:22:02.549 --> 00:22:03.909 the more context you give,
NOTE Confidence: 0.9425882

00:22:03.909 --> 00:22:05.350 the better the response is
NOTE Confidence: 0.9425882

00:22:05.350 --> 00:22:06.230 likely to be. And you
NOTE Confidence: 0.9425882

00:22:06.230 --> 00:22:07.690 can continue to ask additional
NOTE Confidence: 0.917047

00:22:08.005 --> 00:22:09.865 additional questions and eventually,
NOTE Confidence: 0.99099314

00:22:10.725 --> 00:22:11.525 ask it to write a
NOTE Confidence: 0.99099314

00:22:11.525 --> 00:22:13.284 pathology pathology report. I think
NOTE Confidence: 0.99099314

00:22:13.284 --> 00:22:14.005 a lot of people have
NOTE Confidence: 0.99099314

00:22:14.005 --> 00:22:15.365 already seen this demo, so
NOTE Confidence: 0.99099314

00:22:15.365 --> 00:22:16.565 I will I will skip
NOTE Confidence: 0.99099314

00:22:16.565 --> 00:22:17.304 through it.
NOTE Confidence: 0.94478196

00:22:18.005 --> 00:22:19.205 But once it has seen
NOTE Confidence: 0.94478196

00:22:19.205 --> 00:22:20.085 enough, it can it can
NOTE Confidence: 0.94478196

00:22:20.085 --> 00:22:20.965 write up a pathology report.

NOTE Confidence: 0.94478196
00:22:20.965 --> 00:22:21.684 So what we hope is
NOTE Confidence: 0.94478196
00:22:21.684 --> 00:22:22.484 that this would happen in
NOTE Confidence: 0.94478196
00:22:22.484 --> 00:22:23.929 the background where the host
NOTE Confidence: 0.94478196
00:22:23.929 --> 00:22:25.450 level analysis was is already
NOTE Confidence: 0.94478196
00:22:25.450 --> 00:22:26.650 done and the report is
NOTE Confidence: 0.94478196
00:22:26.650 --> 00:22:27.690 already generated by the time
NOTE Confidence: 0.94478196
00:22:27.690 --> 00:22:28.890 a pathologist is looking at
NOTE Confidence: 0.94478196
00:22:28.890 --> 00:22:29.390 it.
NOTE Confidence: 0.9652299
00:22:30.410 --> 00:22:31.850 We're also interested in essentially
NOTE Confidence: 0.9652299
00:22:31.850 --> 00:22:32.890 using this for lower source
NOTE Confidence: 0.9652299
00:22:32.890 --> 00:22:34.490 settings where we have just
NOTE Confidence: 0.9652299
00:22:34.490 --> 00:22:35.770 a cell phone coupled directly
NOTE Confidence: 0.9652299
00:22:35.770 --> 00:22:37.450 to a microscope, taking multiple
NOTE Confidence: 0.9652299
00:22:37.450 --> 00:22:37.950 images
NOTE Confidence: 0.98311216
00:22:38.375 --> 00:22:39.815 from the from the microscope,
NOTE Confidence: 0.98311216

00:22:39.815 --> 00:22:41.655 and then asking questions to
NOTE Confidence: 0.98311216

00:22:41.655 --> 00:22:42.155 the
NOTE Confidence: 0.87066257

00:22:42.455 --> 00:22:43.735 to to the chatbot about
NOTE Confidence: 0.87066257

00:22:43.735 --> 00:22:44.715 what what the,
NOTE Confidence: 0.984095

00:22:45.975 --> 00:22:47.175 what what is essentially in
NOTE Confidence: 0.984095

00:22:47.175 --> 00:22:48.375 those in those images. The
NOTE Confidence: 0.984095

00:22:48.375 --> 00:22:49.815 more context, again, we we
NOTE Confidence: 0.984095

00:22:49.815 --> 00:22:50.920 give it, the better the
NOTE Confidence: 0.984095

00:22:50.920 --> 00:22:52.680 chatbot is likely likely to
NOTE Confidence: 0.984095

00:22:52.680 --> 00:22:54.440 do. So in this case,
NOTE Confidence: 0.984095

00:22:54.440 --> 00:22:55.660 the user is asking
NOTE Confidence: 0.96679026

00:22:56.040 --> 00:22:57.240 that these images are from
NOTE Confidence: 0.96679026

00:22:57.240 --> 00:22:58.359 a patient with a left
NOTE Confidence: 0.96679026

00:22:58.359 --> 00:22:59.560 breast mass. What do you
NOTE Confidence: 0.96679026

00:22:59.560 --> 00:23:00.359 what do you see? And
NOTE Confidence: 0.96679026

00:23:00.359 --> 00:23:01.400 the model would come up

NOTE Confidence: 0.96679026

00:23:01.400 --> 00:23:02.859 with a with a response.

NOTE Confidence: 0.96015555

00:23:03.480 --> 00:23:04.440 And you can continue to

NOTE Confidence: 0.96015555

00:23:04.440 --> 00:23:05.640 chat with the model, ask

NOTE Confidence: 0.96015555

00:23:05.640 --> 00:23:07.635 about, you know, potential treatment

NOTE Confidence: 0.96015555

00:23:07.635 --> 00:23:09.955 guideline or what next steps

NOTE Confidence: 0.96015555

00:23:09.955 --> 00:23:11.415 of the diagnostic process,

NOTE Confidence: 0.95228297

00:23:12.435 --> 00:23:14.115 would this case essentially go

NOTE Confidence: 0.95228297

00:23:14.115 --> 00:23:14.615 through.

NOTE Confidence: 0.9760936

00:23:15.955 --> 00:23:17.155 This example is is quite

NOTE Confidence: 0.9760936

00:23:17.155 --> 00:23:18.455 interesting because we're using

NOTE Confidence: 0.9263333

00:23:19.129 --> 00:23:19.423 coupled to a microscope. And

NOTE Confidence: 0.9263333

00:23:19.423 --> 00:23:20.429 the user is is essentially

NOTE Confidence: 0.998615

00:23:20.889 --> 00:23:21.389 asking

NOTE Confidence: 0.9113606

00:23:22.169 --> 00:23:22.919 what what

NOTE Confidence: 0.9711923

00:23:32.605 --> 00:23:33.484 these images are from a

NOTE Confidence: 0.9711923

00:23:33.484 --> 00:23:34.605 patient with a left breast
NOTE Confidence: 0.9711923

00:23:34.605 --> 00:23:35.565 mass. What what do you
NOTE Confidence: 0.9711923

00:23:35.565 --> 00:23:37.744 see? And, the chatbot
NOTE Confidence: 0.9702021

00:23:38.205 --> 00:23:39.904 responds that, well, this is
NOTE Confidence: 0.6275665

00:23:40.524 --> 00:23:41.024 likely
NOTE Confidence: 0.78240633

00:23:41.644 --> 00:23:42.304 a a a melanoma.
NOTE Confidence: 0.9711717

00:23:42.924 --> 00:23:43.965 And then the user can
NOTE Confidence: 0.9711717

00:23:43.965 --> 00:23:44.460 ask,
NOTE Confidence: 0.99968475

00:23:45.179 --> 00:23:46.940 what ancillary tests could be
NOTE Confidence: 0.99968475

00:23:46.940 --> 00:23:47.440 used
NOTE Confidence: 0.9909102

00:23:47.820 --> 00:23:48.480 to confirm,
NOTE Confidence: 0.9747211

00:23:49.100 --> 00:23:50.720 this this particular case.
NOTE Confidence: 0.96807665

00:23:51.740 --> 00:23:53.020 So what what IHC should
NOTE Confidence: 0.96807665

00:23:53.020 --> 00:23:54.220 be, should I order to
NOTE Confidence: 0.96807665

00:23:54.220 --> 00:23:55.119 confirm this?
NOTE Confidence: 0.96379036

00:23:55.500 --> 00:23:56.000 And,

NOTE Confidence: 0.9367937
00:23:56.619 --> 00:23:57.679 it can give suggestions
NOTE Confidence: 0.99823123
00:23:58.059 --> 00:23:58.880 for the IHCs
NOTE Confidence: 0.95775557
00:23:59.260 --> 00:24:00.485 that that can be used
NOTE Confidence: 0.95775557
00:24:00.485 --> 00:24:02.005 to confirm this. Once those
NOTE Confidence: 0.95775557
00:24:02.005 --> 00:24:04.405 ancillary tests are in, you
NOTE Confidence: 0.95775557
00:24:04.405 --> 00:24:06.085 can image those ancillary the
NOTE Confidence: 0.95775557
00:24:06.085 --> 00:24:07.525 the those slides again with
NOTE Confidence: 0.95775557
00:24:07.525 --> 00:24:08.745 your with your microscope
NOTE Confidence: 0.6324951
00:24:09.365 --> 00:24:09.865 and,
NOTE Confidence: 0.9175595
00:24:11.445 --> 00:24:12.725 you skip that slide within
NOTE Confidence: 0.9175595
00:24:12.725 --> 00:24:13.869 the same context. So this
NOTE Confidence: 0.9175595
00:24:13.869 --> 00:24:14.909 is this is where the
NOTE Confidence: 0.9175595
00:24:14.909 --> 00:24:15.950 sort of the innovation is
NOTE Confidence: 0.9175595
00:24:15.950 --> 00:24:16.830 is is that you can
NOTE Confidence: 0.9175595
00:24:16.830 --> 00:24:18.830 continue to ask questions within
NOTE Confidence: 0.9175595

00:24:18.830 --> 00:24:19.730 the same context.
NOTE Confidence: 0.91373426

00:24:20.109 --> 00:24:21.309 So, essentially, what this means
NOTE Confidence: 0.91373426

00:24:21.309 --> 00:24:22.109 is is that at a
NOTE Confidence: 0.91373426

00:24:22.109 --> 00:24:23.309 whole slide level, if you
NOTE Confidence: 0.91373426

00:24:23.309 --> 00:24:24.929 have an h HNE,
NOTE Confidence: 0.95959616

00:24:25.654 --> 00:24:26.774 the the model can predict
NOTE Confidence: 0.95959616

00:24:26.774 --> 00:24:28.615 what ancillary which ancillary test
NOTE Confidence: 0.95959616

00:24:28.615 --> 00:24:29.414 needs to need to be
NOTE Confidence: 0.95959616

00:24:29.414 --> 00:24:31.095 ordered, order those tests on
NOTE Confidence: 0.95959616

00:24:31.095 --> 00:24:32.855 its own, and then ingest
NOTE Confidence: 0.95959616

00:24:32.855 --> 00:24:34.054 those images in the same
NOTE Confidence: 0.95959616

00:24:34.054 --> 00:24:35.335 context and have a pathology
NOTE Confidence: 0.95959616

00:24:35.335 --> 00:24:36.734 report ready by a time
NOTE Confidence: 0.95959616

00:24:36.734 --> 00:24:37.514 a pathologist
NOTE Confidence: 0.95898974

00:24:37.894 --> 00:24:38.934 essentially starts looking at it.
NOTE Confidence: 0.95898974

00:24:38.934 --> 00:24:39.575 So that's what we are

NOTE Confidence: 0.95898974
00:24:39.575 --> 00:24:41.160 working towards. But with the
NOTE Confidence: 0.95898974
00:24:41.160 --> 00:24:42.360 with the model, obviously, we
NOTE Confidence: 0.95898974
00:24:42.360 --> 00:24:43.640 needed to do a lot
NOTE Confidence: 0.95898974
00:24:43.640 --> 00:24:45.480 of evaluation. So this was
NOTE Confidence: 0.95898974
00:24:45.480 --> 00:24:46.760 done in sort of a
NOTE Confidence: 0.95898974
00:24:46.760 --> 00:24:48.360 two two pronged strategy with
NOTE Confidence: 0.95898974
00:24:48.360 --> 00:24:49.820 a lot of quantitative evaluation
NOTE Confidence: 0.95898974
00:24:50.040 --> 00:24:50.540 with,
NOTE Confidence: 0.96370655
00:24:51.080 --> 00:24:52.520 multiple choice questions and how
NOTE Confidence: 0.96370655
00:24:52.520 --> 00:24:53.560 well the model does in
NOTE Confidence: 0.96370655
00:24:53.560 --> 00:24:54.600 comparison to some of the
NOTE Confidence: 0.96370655
00:24:54.600 --> 00:24:55.800 other models including g p
NOTE Confidence: 0.96370655
00:24:55.800 --> 00:24:57.105 d four o or generically
NOTE Confidence: 0.96370655
00:24:57.105 --> 00:24:58.305 trained models and and and
NOTE Confidence: 0.96370655
00:24:58.305 --> 00:24:59.665 models that are trained specifically
NOTE Confidence: 0.96370655

00:24:59.665 --> 00:25:00.725 on medical data.
NOTE Confidence: 0.9463008

00:25:01.505 --> 00:25:02.405 But then also,
NOTE Confidence: 0.979539

00:25:03.425 --> 00:25:05.185 by comparing it with, with
NOTE Confidence: 0.979539

00:25:05.185 --> 00:25:06.945 pathologists and asking pathologists to
NOTE Confidence: 0.979539

00:25:06.945 --> 00:25:08.725 rank how well the response
NOTE Confidence: 0.979539

00:25:08.785 --> 00:25:09.285 is
NOTE Confidence: 0.95112115

00:25:09.700 --> 00:25:10.420 or or or how well
NOTE Confidence: 0.95112115

00:25:10.420 --> 00:25:11.460 the monitor does in terms
NOTE Confidence: 0.95112115

00:25:11.460 --> 00:25:12.820 of the response in comparison
NOTE Confidence: 0.95112115

00:25:12.820 --> 00:25:14.100 to some of the other
NOTE Confidence: 0.95112115

00:25:14.100 --> 00:25:15.380 other models. And we overall,
NOTE Confidence: 0.95112115

00:25:15.380 --> 00:25:16.340 we see that this model
NOTE Confidence: 0.95112115

00:25:16.340 --> 00:25:18.020 that's specifically trained on lots
NOTE Confidence: 0.95112115

00:25:18.020 --> 00:25:19.700 of pathology data does quite
NOTE Confidence: 0.95112115

00:25:19.700 --> 00:25:21.640 well on diagnosis microscopy,
NOTE Confidence: 0.9645629

00:25:22.775 --> 00:25:24.054 based questions, but does not

NOTE Confidence: 0.9645629
00:25:24.054 --> 00:25:25.575 do very well on on
NOTE Confidence: 0.9645629
00:25:25.575 --> 00:25:27.015 clinical questions. And the reason
NOTE Confidence: 0.9645629
00:25:27.015 --> 00:25:28.054 largely is that we did
NOTE Confidence: 0.9645629
00:25:28.054 --> 00:25:29.174 not train with lots of
NOTE Confidence: 0.9645629
00:25:29.174 --> 00:25:30.375 lots and lots of medical
NOTE Confidence: 0.9645629
00:25:30.375 --> 00:25:31.815 texts that that GPD four
NOTE Confidence: 0.9645629
00:25:31.815 --> 00:25:33.755 o is largely trained on.
NOTE Confidence: 0.98735887
00:25:34.695 --> 00:25:35.595 So I've seen
NOTE Confidence: 0.9754819
00:25:35.920 --> 00:25:36.960 that you can build lots
NOTE Confidence: 0.9754819
00:25:36.960 --> 00:25:38.960 of supervised models, self supervised
NOTE Confidence: 0.9754819
00:25:38.960 --> 00:25:40.260 models for feature extraction,
NOTE Confidence: 0.9872233
00:25:40.800 --> 00:25:42.320 contrast with other modalities to
NOTE Confidence: 0.9872233
00:25:42.320 --> 00:25:43.300 get richer feature
NOTE Confidence: 0.9814934
00:25:43.760 --> 00:25:45.760 richer features, and then use
NOTE Confidence: 0.9814934
00:25:45.760 --> 00:25:47.060 those features in a generative
NOTE Confidence: 0.9814934

00:25:47.119 --> 00:25:48.240 AI setting where you have
NOTE Confidence: 0.9814934

00:25:48.240 --> 00:25:49.359 a singular model that can
NOTE Confidence: 0.9814934

00:25:49.359 --> 00:25:50.180 harbor information
NOTE Confidence: 0.9811181

00:25:50.744 --> 00:25:52.265 ideally around all of human
NOTE Confidence: 0.9811181

00:25:52.265 --> 00:25:53.625 pathology or eventually around all
NOTE Confidence: 0.9811181

00:25:53.625 --> 00:25:54.525 of human pathology.
NOTE Confidence: 0.9961735

00:25:54.905 --> 00:25:56.105 But what can you do
NOTE Confidence: 0.9961735

00:25:56.105 --> 00:25:57.465 with the generative AI and
NOTE Confidence: 0.9961735

00:25:57.465 --> 00:25:58.845 all these features? So
NOTE Confidence: 0.96911573

00:25:59.145 --> 00:26:00.345 the the next obvious step
NOTE Confidence: 0.96911573

00:26:00.345 --> 00:26:01.145 is that you can build
NOTE Confidence: 0.96911573

00:26:01.145 --> 00:26:01.945 an agent on top of
NOTE Confidence: 0.96911573

00:26:01.945 --> 00:26:02.445 it.
NOTE Confidence: 0.97173464

00:26:03.320 --> 00:26:04.200 So you have a bunch
NOTE Confidence: 0.97173464

00:26:04.200 --> 00:26:05.320 of what if questions and
NOTE Confidence: 0.97173464

00:26:05.320 --> 00:26:06.200 about and and what if

NOTE Confidence: 0.97173464
00:26:06.200 --> 00:26:07.560 statements. Right? So what what
NOTE Confidence: 0.97173464
00:26:07.560 --> 00:26:08.600 if agents could do all
NOTE Confidence: 0.97173464
00:26:08.600 --> 00:26:10.119 the biomedical data analysis for
NOTE Confidence: 0.97173464
00:26:10.119 --> 00:26:11.240 you? What if AI agents
NOTE Confidence: 0.97173464
00:26:11.240 --> 00:26:13.260 could develop, assess, and explain
NOTE Confidence: 0.97173464
00:26:13.400 --> 00:26:15.500 AI models for for pathology?
NOTE Confidence: 0.9902375
00:26:16.065 --> 00:26:17.264 What if an AI agent
NOTE Confidence: 0.9902375
00:26:17.264 --> 00:26:18.804 could write code, run experiments,
NOTE Confidence: 0.9902375
00:26:18.865 --> 00:26:19.605 test hypotheses?
NOTE Confidence: 0.9818652
00:26:20.225 --> 00:26:21.184 What if an AI agent
NOTE Confidence: 0.9818652
00:26:21.184 --> 00:26:23.205 could continuously run-in the background
NOTE Confidence: 0.9818652
00:26:23.264 --> 00:26:25.125 attempting to find common morphologic
NOTE Confidence: 0.9818652
00:26:25.264 --> 00:26:27.264 features across patient cohorts and
NOTE Confidence: 0.9818652
00:26:27.264 --> 00:26:28.830 correlate with outcome? Right? So
NOTE Confidence: 0.9818652
00:26:28.830 --> 00:26:29.570 so these are
NOTE Confidence: 0.9331401

00:26:30.190 --> 00:26:31.549 all questions we have because
NOTE Confidence: 0.9331401

00:26:31.549 --> 00:26:32.929 we we were living in
NOTE Confidence: 0.9331401

00:26:33.150 --> 00:26:33.950 in in sort of the
NOTE Confidence: 0.9331401

00:26:33.950 --> 00:26:35.970 age of AI agents where,
NOTE Confidence: 0.92745423

00:26:36.669 --> 00:26:37.869 these agents would do things
NOTE Confidence: 0.92745423

00:26:37.869 --> 00:26:38.990 for us. So that's what
NOTE Confidence: 0.92745423

00:26:38.990 --> 00:26:40.029 we essentially try to do.
NOTE Confidence: 0.92745423

00:26:40.029 --> 00:26:40.990 We build we we built
NOTE Confidence: 0.92745423

00:26:40.990 --> 00:26:41.950 the agent on top of,
NOTE Confidence: 0.9459208

00:26:43.555 --> 00:26:45.234 the the generative AI model
NOTE Confidence: 0.9459208

00:26:45.234 --> 00:26:45.975 that we had.
NOTE Confidence: 0.95851326

00:26:46.675 --> 00:26:47.875 It uses all the code
NOTE Confidence: 0.95851326

00:26:47.875 --> 00:26:48.835 that we had developed over
NOTE Confidence: 0.95851326

00:26:48.835 --> 00:26:49.895 the past five years,
NOTE Confidence: 0.9591151

00:26:50.515 --> 00:26:51.635 but it's capable of writing
NOTE Confidence: 0.9591151

00:26:51.635 --> 00:26:53.075 additional code to patch through

NOTE Confidence: 0.9591151
00:26:53.075 --> 00:26:53.555 multiple,
NOTE Confidence: 0.95200473
00:26:55.220 --> 00:26:56.340 parts of the of of
NOTE Confidence: 0.95200473
00:26:56.340 --> 00:26:57.460 the library that we that
NOTE Confidence: 0.95200473
00:26:57.460 --> 00:26:58.580 we had. I'll show a
NOTE Confidence: 0.95200473
00:26:58.580 --> 00:27:00.660 few example use cases. So
NOTE Confidence: 0.95200473
00:27:00.660 --> 00:27:01.780 in this in this particular
NOTE Confidence: 0.95200473
00:27:01.780 --> 00:27:03.140 use case, the user is
NOTE Confidence: 0.95200473
00:27:03.140 --> 00:27:03.960 saying that,
NOTE Confidence: 0.94627655
00:27:05.060 --> 00:27:06.100 can you help me train
NOTE Confidence: 0.94627655
00:27:06.100 --> 00:27:07.619 a model to classify classify
NOTE Confidence: 0.94627655
00:27:07.619 --> 00:27:09.115 between responders and and non
NOTE Confidence: 0.94627655
00:27:09.115 --> 00:27:10.395 responders? And here's where my
NOTE Confidence: 0.94627655
00:27:10.395 --> 00:27:11.674 data is stored. And this
NOTE Confidence: 0.94627655
00:27:11.674 --> 00:27:12.794 is what the magnification it
NOTE Confidence: 0.94627655
00:27:12.794 --> 00:27:14.335 was it was trained on.
NOTE Confidence: 0.89553666

00:27:15.115 --> 00:27:16.395 So the journey of AI
NOTE Confidence: 0.89553666

00:27:16.395 --> 00:27:17.355 aspect of this is that
NOTE Confidence: 0.89553666

00:27:17.355 --> 00:27:18.315 is it it it can
NOTE Confidence: 0.89553666

00:27:18.315 --> 00:27:19.914 generate a plan. And the
NOTE Confidence: 0.89553666

00:27:19.914 --> 00:27:21.775 plan can then retrieve and,
NOTE Confidence: 0.9377807

00:27:22.234 --> 00:27:23.595 use code that's already written
NOTE Confidence: 0.9377807

00:27:23.595 --> 00:27:24.690 as well as writing some
NOTE Confidence: 0.9377807

00:27:24.690 --> 00:27:26.049 additional code that might be
NOTE Confidence: 0.9377807

00:27:26.049 --> 00:27:27.409 needed to pass through the
NOTE Confidence: 0.9377807

00:27:27.570 --> 00:27:28.450 what what code is already
NOTE Confidence: 0.9377807

00:27:28.450 --> 00:27:28.950 available.
NOTE Confidence: 0.96102005

00:27:29.250 --> 00:27:30.230 So in in in computational
NOTE Confidence: 0.96102005

00:27:30.369 --> 00:27:31.830 biology and largely in bioinformatics,
NOTE Confidence: 0.96102005

00:27:32.049 --> 00:27:33.750 we're often using large libraries,
NOTE Confidence: 0.9761758

00:27:34.369 --> 00:27:36.210 existing existing code to analyze
NOTE Confidence: 0.9761758

00:27:36.210 --> 00:27:37.490 new kinds of data leading

NOTE Confidence: 0.9761758

00:27:37.490 --> 00:27:38.629 to leading to discovery.

NOTE Confidence: 0.97790366

00:27:39.994 --> 00:27:40.655 And there's

NOTE Confidence: 0.9809349

00:27:40.955 --> 00:27:41.994 some code writing, but it's

NOTE Confidence: 0.9809349

00:27:41.994 --> 00:27:43.755 often patching through code that's

NOTE Confidence: 0.9809349

00:27:43.755 --> 00:27:45.595 already been been written. And

NOTE Confidence: 0.9809349

00:27:45.595 --> 00:27:47.195 that's what's essentially happening on

NOTE Confidence: 0.9809349

00:27:47.195 --> 00:27:48.475 its own here. So so

NOTE Confidence: 0.9809349

00:27:48.475 --> 00:27:49.215 we use,

NOTE Confidence: 0.8489271

00:27:50.234 --> 00:27:51.355 path chat to as a

NOTE Confidence: 0.8489271

00:27:51.355 --> 00:27:52.415 morphologic descriptor,

NOTE Confidence: 0.9461207

00:27:53.440 --> 00:27:55.200 and existing code that that's

NOTE Confidence: 0.9461207

00:27:55.200 --> 00:27:56.080 used in retrieve. So the

NOTE Confidence: 0.9461207

00:27:56.080 --> 00:27:57.119 model says that, well, here's

NOTE Confidence: 0.9461207

00:27:57.119 --> 00:27:58.260 an ROC curve.

NOTE Confidence: 0.9616393

00:27:58.800 --> 00:27:59.600 It used about a hundred

NOTE Confidence: 0.9616393

00:27:59.600 --> 00:28:00.640 cases to train the model,
NOTE Confidence: 0.9616393

00:28:00.640 --> 00:28:01.680 and the ROC is zero
NOTE Confidence: 0.9616393

00:28:01.680 --> 00:28:02.640 point eight five five. And
NOTE Confidence: 0.9616393

00:28:02.640 --> 00:28:03.920 the next question can be
NOTE Confidence: 0.9616393

00:28:03.920 --> 00:28:04.720 that, well, can I look
NOTE Confidence: 0.9616393

00:28:04.720 --> 00:28:06.480 at a heat map showing
NOTE Confidence: 0.9616393

00:28:06.480 --> 00:28:07.600 what the high tension regions
NOTE Confidence: 0.9616393

00:28:07.600 --> 00:28:08.085 are,
NOTE Confidence: 0.9384917

00:28:09.684 --> 00:28:10.565 and and what what the
NOTE Confidence: 0.9384917

00:28:10.565 --> 00:28:11.524 model is using to make
NOTE Confidence: 0.9384917

00:28:11.524 --> 00:28:12.825 these classification determinations?
NOTE Confidence: 0.963764

00:28:13.924 --> 00:28:14.725 So you can you can
NOTE Confidence: 0.963764

00:28:14.725 --> 00:28:16.005 do that. And then we
NOTE Confidence: 0.963764

00:28:16.005 --> 00:28:17.524 can invoke past chat or
NOTE Confidence: 0.963764

00:28:17.524 --> 00:28:19.044 the generative AI model to
NOTE Confidence: 0.963764

00:28:19.044 --> 00:28:20.484 write a report about what

NOTE Confidence: 0.963764
00:28:20.484 --> 00:28:21.684 the model used in making
NOTE Confidence: 0.963764
00:28:21.684 --> 00:28:22.984 these classification determinations.
NOTE Confidence: 0.9074862
00:28:23.609 --> 00:28:24.650 What were the high tension
NOTE Confidence: 0.9074862
00:28:24.650 --> 00:28:25.609 regions? What were the low
NOTE Confidence: 0.9074862
00:28:25.609 --> 00:28:26.430 tension regions?
NOTE Confidence: 0.897472
00:28:27.609 --> 00:28:28.250 And it can do that.
NOTE Confidence: 0.897472
00:28:28.250 --> 00:28:29.369 So it so basically says
NOTE Confidence: 0.897472
00:28:29.369 --> 00:28:30.890 that the the model used
NOTE Confidence: 0.897472
00:28:30.890 --> 00:28:32.910 inflammatory regions, necrosis, and fibrosis
NOTE Confidence: 0.95160824
00:28:33.690 --> 00:28:35.369 to determine which cohort is
NOTE Confidence: 0.95160824
00:28:35.369 --> 00:28:36.730 the responder versus which which
NOTE Confidence: 0.95160824
00:28:36.730 --> 00:28:38.010 patient is the responder versus
NOTE Confidence: 0.95160824
00:28:38.010 --> 00:28:39.375 nonresponder. Responder. And then you
NOTE Confidence: 0.95160824
00:28:39.375 --> 00:28:40.335 might wanna do some more
NOTE Confidence: 0.95160824
00:28:40.335 --> 00:28:42.015 fine grained analysis, like, segment
NOTE Confidence: 0.95160824

00:28:42.015 --> 00:28:43.535 all the cells, classify them,
NOTE Confidence: 0.95160824

00:28:43.535 --> 00:28:45.055 and then run handcrafted feature
NOTE Confidence: 0.95160824

00:28:45.055 --> 00:28:46.115 analysis on top,
NOTE Confidence: 0.9932961

00:28:46.415 --> 00:28:47.215 and it would do that
NOTE Confidence: 0.9932961

00:28:47.215 --> 00:28:48.655 for you, for you as
NOTE Confidence: 0.9932961

00:28:48.655 --> 00:28:49.635 well. So
NOTE Confidence: 0.99831057

00:28:50.015 --> 00:28:50.515 the
NOTE Confidence: 0.9822985

00:28:50.895 --> 00:28:51.395 the,
NOTE Confidence: 0.9811781

00:28:52.335 --> 00:28:54.050 what's essentially being done is
NOTE Confidence: 0.9811781

00:28:54.050 --> 00:28:54.550 that
NOTE Confidence: 0.9341061

00:28:54.930 --> 00:28:56.370 the generative AI can parse
NOTE Confidence: 0.9341061

00:28:56.370 --> 00:28:58.050 your text command, convert it
NOTE Confidence: 0.9341061

00:28:58.050 --> 00:28:59.410 into or oppose it as
NOTE Confidence: 0.9341061

00:28:59.410 --> 00:29:00.550 a machine learning problem,
NOTE Confidence: 0.9989323

00:29:01.010 --> 00:29:01.510 then
NOTE Confidence: 0.9977895

00:29:01.970 --> 00:29:02.470 recall

NOTE Confidence: 0.968838
00:29:02.770 --> 00:29:03.650 all the code that was
NOTE Confidence: 0.968838
00:29:03.650 --> 00:29:05.170 written potentially for other purposes
NOTE Confidence: 0.968838
00:29:05.170 --> 00:29:06.390 but has been now streamlined,
NOTE Confidence: 0.968838
00:29:06.605 --> 00:29:07.725 and write additional code where
NOTE Confidence: 0.968838
00:29:07.725 --> 00:29:08.705 it needs to be written,
NOTE Confidence: 0.98338675
00:29:09.325 --> 00:29:10.925 to patch everything everything together.
NOTE Confidence: 0.98338675
00:29:10.925 --> 00:29:12.305 Another example of this
NOTE Confidence: 0.974951
00:29:13.005 --> 00:29:15.085 is around case retrieval. So
NOTE Confidence: 0.974951
00:29:15.085 --> 00:29:15.965 the user is saying that
NOTE Confidence: 0.974951
00:29:15.965 --> 00:29:16.925 can you help me build
NOTE Confidence: 0.974951
00:29:16.925 --> 00:29:18.525 a database of slides that
NOTE Confidence: 0.974951
00:29:18.525 --> 00:29:20.065 I can use to query,
NOTE Confidence: 0.83425146
00:29:21.409 --> 00:29:22.309 query my databases
NOTE Confidence: 0.9148376
00:29:22.690 --> 00:29:23.970 or or query my large
NOTE Confidence: 0.9148376
00:29:23.970 --> 00:29:25.169 database of of all host
NOTE Confidence: 0.9148376

00:29:25.169 --> 00:29:26.529 side images. So it uses
NOTE Confidence: 0.9148376

00:29:26.529 --> 00:29:27.730 the whole slide level foundation
NOTE Confidence: 0.9148376

00:29:27.730 --> 00:29:28.690 model that we built to
NOTE Confidence: 0.9148376

00:29:28.690 --> 00:29:30.149 extract a single feature representation
NOTE Confidence: 0.9148376

00:29:30.289 --> 00:29:30.789 corresponding
NOTE Confidence: 0.97844917

00:29:31.090 --> 00:29:32.769 each image and builds the
NOTE Confidence: 0.97844917

00:29:32.769 --> 00:29:33.269 entire
NOTE Confidence: 0.9925238

00:29:33.595 --> 00:29:35.535 entire database. Once the database
NOTE Confidence: 0.9925238

00:29:35.675 --> 00:29:36.875 is is built, the user
NOTE Confidence: 0.9925238

00:29:36.875 --> 00:29:38.235 can give a single slide
NOTE Confidence: 0.9925238

00:29:38.235 --> 00:29:39.275 and say, can you find
NOTE Confidence: 0.9925238

00:29:39.275 --> 00:29:40.095 common cases
NOTE Confidence: 0.95051503

00:29:40.715 --> 00:29:41.915 the most common cases to
NOTE Confidence: 0.95051503

00:29:41.915 --> 00:29:42.815 this this particular
NOTE Confidence: 0.9260877

00:29:43.115 --> 00:29:44.175 this particular image?
NOTE Confidence: 0.93248457

00:29:45.275 --> 00:29:46.940 So retrieve the the the

NOTE Confidence: 0.93248457
00:29:46.940 --> 00:29:48.060 the top three most similar
NOTE Confidence: 0.93248457
00:29:48.060 --> 00:29:49.420 cases to this this particular
NOTE Confidence: 0.93248457
00:29:49.420 --> 00:29:50.780 image. And this these images
NOTE Confidence: 0.93248457
00:29:50.780 --> 00:29:51.440 are from,
NOTE Confidence: 0.69248796
00:29:52.020 --> 00:29:53.600 I believe, leiomyosarcoma.
NOTE Confidence: 0.97395325
00:29:54.380 --> 00:29:54.880 So,
NOTE Confidence: 0.9889135
00:29:55.260 --> 00:29:56.300 it would retrieve the three
NOTE Confidence: 0.9889135
00:29:56.300 --> 00:29:58.160 most common images. This is
NOTE Confidence: 0.9889135
00:29:58.220 --> 00:29:59.120 based on the TCGA.
NOTE Confidence: 0.9872394
00:30:00.735 --> 00:30:01.934 And then you can just
NOTE Confidence: 0.9872394
00:30:01.934 --> 00:30:02.915 introspect them.
NOTE Confidence: 0.94961894
00:30:04.575 --> 00:30:05.075 And
NOTE Confidence: 0.96618927
00:30:05.375 --> 00:30:06.995 the the the last example
NOTE Confidence: 0.9594148
00:30:07.615 --> 00:30:09.615 is around multimodal data data
NOTE Confidence: 0.9594148
00:30:09.615 --> 00:30:11.215 integration. So the user can
NOTE Confidence: 0.9594148

00:30:11.215 --> 00:30:12.415 say that, well, train a
NOTE Confidence: 0.9594148

00:30:12.415 --> 00:30:13.635 multimodal data integration,
NOTE Confidence: 0.90593076

00:30:14.495 --> 00:30:16.174 model using something very basic
NOTE Confidence: 0.90593076

00:30:16.174 --> 00:30:17.090 like the concat
NOTE Confidence: 0.9724239

00:30:17.390 --> 00:30:19.630 functionality. So, basically, the way
NOTE Confidence: 0.9724239

00:30:19.630 --> 00:30:21.150 this works or or how
NOTE Confidence: 0.9724239

00:30:21.150 --> 00:30:22.830 our training data essentially work
NOTE Confidence: 0.9724239

00:30:22.830 --> 00:30:24.370 is is that if you
NOTE Confidence: 0.9724239

00:30:24.590 --> 00:30:25.789 the more specific you make
NOTE Confidence: 0.9724239

00:30:25.789 --> 00:30:27.150 your prompt, it will use
NOTE Confidence: 0.9724239

00:30:27.150 --> 00:30:28.270 all the information from your
NOTE Confidence: 0.9724239

00:30:28.270 --> 00:30:29.865 prompt. If it's if if
NOTE Confidence: 0.9724239

00:30:29.865 --> 00:30:30.985 that information is not there,
NOTE Confidence: 0.9724239

00:30:30.985 --> 00:30:32.365 it would make some assumptions.
NOTE Confidence: 0.96012014

00:30:32.745 --> 00:30:33.865 And if it can't make
NOTE Confidence: 0.96012014

00:30:33.865 --> 00:30:35.145 make assumptions, for example, where

NOTE Confidence: 0.96012014
00:30:35.145 --> 00:30:36.185 your data is located or
NOTE Confidence: 0.96012014
00:30:36.185 --> 00:30:37.465 how your data is is
NOTE Confidence: 0.96012014
00:30:37.465 --> 00:30:38.825 organized, it will ask additional
NOTE Confidence: 0.96012014
00:30:38.825 --> 00:30:39.325 questions.
NOTE Confidence: 0.9852532
00:30:39.705 --> 00:30:40.205 So,
NOTE Confidence: 0.97117984
00:30:41.065 --> 00:30:42.265 in this particular case, it's
NOTE Confidence: 0.97117984
00:30:42.265 --> 00:30:42.985 training a model,
NOTE Confidence: 0.986304
00:30:44.539 --> 00:30:46.320 to separate low grade glioma
NOTE Confidence: 0.986304
00:30:46.380 --> 00:30:48.159 cases by integrating histology,
NOTE Confidence: 0.96746606
00:30:48.539 --> 00:30:50.940 molecular, and radiology data. So
NOTE Confidence: 0.96746606
00:30:50.940 --> 00:30:53.179 it extract features, integrates them,
NOTE Confidence: 0.96746606
00:30:53.179 --> 00:30:54.480 and comes up with this
NOTE Confidence: 0.96746606
00:30:54.539 --> 00:30:56.299 with this, Kilometers curve showing
NOTE Confidence: 0.96746606
00:30:56.299 --> 00:30:57.100 that, well, this is how
NOTE Confidence: 0.96746606
00:30:57.100 --> 00:30:57.600 separable
NOTE Confidence: 0.981837

00:30:58.059 --> 00:31:00.065 the the patients are. The
NOTE Confidence: 0.981837

00:31:00.065 --> 00:31:01.585 next question is that that
NOTE Confidence: 0.981837

00:31:01.585 --> 00:31:03.365 can you help me introspect
NOTE Confidence: 0.981837

00:31:03.505 --> 00:31:04.725 the the genomics?
NOTE Confidence: 0.9869901

00:31:05.184 --> 00:31:06.785 What molecular features is the
NOTE Confidence: 0.9869901

00:31:06.785 --> 00:31:08.385 model using in making these
NOTE Confidence: 0.9869901

00:31:08.385 --> 00:31:09.445 making these determinations?
NOTE Confidence: 0.93652654

00:31:10.225 --> 00:31:11.910 So it it plots that
NOTE Confidence: 0.93652654

00:31:12.150 --> 00:31:13.190 and continue to ask more
NOTE Confidence: 0.93652654

00:31:13.190 --> 00:31:14.390 questions. Can I look at
NOTE Confidence: 0.93652654

00:31:14.390 --> 00:31:14.890 the,
NOTE Confidence: 0.95598716

00:31:15.910 --> 00:31:16.870 the heat map, or can
NOTE Confidence: 0.95598716

00:31:16.870 --> 00:31:17.850 I look at the radiology,
NOTE Confidence: 0.9997559

00:31:19.510 --> 00:31:20.010 image
NOTE Confidence: 0.9058569

00:31:20.310 --> 00:31:21.830 and what was most important
NOTE Confidence: 0.9058569

00:31:21.830 --> 00:31:23.270 in the radiology image? And

NOTE Confidence: 0.9058569
00:31:23.270 --> 00:31:24.570 it would would essentially,
NOTE Confidence: 0.9338023
00:31:25.510 --> 00:31:26.550 give you a heat map
NOTE Confidence: 0.9338023
00:31:26.550 --> 00:31:27.690 of the of radiology
NOTE Confidence: 0.976784
00:31:29.485 --> 00:31:30.385 image. So
NOTE Confidence: 0.9372929
00:31:31.085 --> 00:31:31.805 it it can do the
NOTE Confidence: 0.9372929
00:31:31.805 --> 00:31:33.405 same for for pathology. So
NOTE Confidence: 0.9372929
00:31:33.405 --> 00:31:34.145 I'll I'll
NOTE Confidence: 0.9765289
00:31:35.165 --> 00:31:36.685 skip through in the interest
NOTE Confidence: 0.9765289
00:31:36.685 --> 00:31:38.045 of time, but, there's there's
NOTE Confidence: 0.9765289
00:31:38.045 --> 00:31:39.825 more information available about this,
NOTE Confidence: 0.9779573
00:31:40.765 --> 00:31:41.965 and we will hopefully put
NOTE Confidence: 0.9779573
00:31:41.965 --> 00:31:43.485 a preprint around the agent
NOTE Confidence: 0.9779573
00:31:43.485 --> 00:31:43.800 out,
NOTE Confidence: 0.9874387
00:31:44.520 --> 00:31:45.660 soon. It's it's,
NOTE Confidence: 0.98775625
00:31:45.960 --> 00:31:47.320 it's it's been in the
NOTE Confidence: 0.98775625

00:31:47.320 --> 00:31:48.440 works for about a year
NOTE Confidence: 0.98775625

00:31:48.440 --> 00:31:50.040 and a half, but, it's
NOTE Confidence: 0.98775625

00:31:50.040 --> 00:31:51.320 been a challenge to do
NOTE Confidence: 0.98775625

00:31:51.320 --> 00:31:53.020 all the evaluations around this.
NOTE Confidence: 0.98043203

00:31:54.680 --> 00:31:55.720 The next thing I wanna
NOTE Confidence: 0.98043203

00:31:55.720 --> 00:31:57.225 talk about is is transitioning
NOTE Confidence: 0.98043203

00:31:57.225 --> 00:31:58.745 from, two d to three
NOTE Confidence: 0.98043203

00:31:58.745 --> 00:31:59.945 d pathology. So I think
NOTE Confidence: 0.98043203

00:31:59.945 --> 00:32:00.445 that
NOTE Confidence: 0.99190164

00:32:00.905 --> 00:32:02.905 everyone here would would acknowledge
NOTE Confidence: 0.99190164

00:32:02.905 --> 00:32:03.705 that we need to have
NOTE Confidence: 0.99190164

00:32:03.705 --> 00:32:04.505 some form of three d
NOTE Confidence: 0.99190164

00:32:04.505 --> 00:32:05.485 pathology because
NOTE Confidence: 0.9496974

00:32:05.785 --> 00:32:06.905 the the the tissue we
NOTE Confidence: 0.9496974

00:32:06.905 --> 00:32:07.865 looked at look at is
NOTE Confidence: 0.9496974

00:32:07.865 --> 00:32:09.420 a very small sample of

NOTE Confidence: 0.9496974
00:32:09.420 --> 00:32:10.160 the actual
NOTE Confidence: 0.97866
00:32:10.620 --> 00:32:11.580 three-dimensional tissue. And there have
NOTE Confidence: 0.97866
00:32:11.580 --> 00:32:12.940 been multitude of studies showing
NOTE Confidence: 0.97866
00:32:12.940 --> 00:32:13.660 that if you look at,
NOTE Confidence: 0.97866
00:32:13.660 --> 00:32:14.880 you know, multiple sections,
NOTE Confidence: 0.9112431
00:32:15.340 --> 00:32:16.240 entire volume,
NOTE Confidence: 0.9726131
00:32:16.620 --> 00:32:18.000 the diagnosis can change,
NOTE Confidence: 0.9286219
00:32:19.100 --> 00:32:20.220 and so forth. And there
NOTE Confidence: 0.9286219
00:32:20.220 --> 00:32:21.100 are a number of different
NOTE Confidence: 0.9286219
00:32:21.100 --> 00:32:22.540 technologies available for this now.
NOTE Confidence: 0.9286219
00:32:22.540 --> 00:32:23.615 So we have OTLS, OTLS,
NOTE Confidence: 0.9286219
00:32:23.615 --> 00:32:24.275 micro CT,
NOTE Confidence: 0.9649806
00:32:24.735 --> 00:32:26.495 as well as, newer techniques
NOTE Confidence: 0.9649806
00:32:26.495 --> 00:32:27.215 where you can take lots
NOTE Confidence: 0.9649806
00:32:27.215 --> 00:32:28.255 of sections and use machine
NOTE Confidence: 0.9649806

00:32:28.255 --> 00:32:29.775 learning to reconstruct the tissue
NOTE Confidence: 0.9649806

00:32:29.775 --> 00:32:30.835 tissue like coda.
NOTE Confidence: 0.958937

00:32:33.135 --> 00:32:34.095 A key issue in the
NOTE Confidence: 0.958937

00:32:34.095 --> 00:32:35.775 adoption of these technologies is
NOTE Confidence: 0.958937

00:32:35.775 --> 00:32:36.835 that how would a pathologist
NOTE Confidence: 0.958937

00:32:37.055 --> 00:32:38.290 look at such a large
NOTE Confidence: 0.958937

00:32:38.290 --> 00:32:39.330 volume? It would take a
NOTE Confidence: 0.958937

00:32:39.330 --> 00:32:39.830 substantially,
NOTE Confidence: 0.95927864

00:32:40.930 --> 00:32:42.390 large amount of time to
NOTE Confidence: 0.95927864

00:32:42.530 --> 00:32:43.410 to to look at each
NOTE Confidence: 0.95927864

00:32:43.410 --> 00:32:44.310 one of these volumes,
NOTE Confidence: 0.9770594

00:32:44.850 --> 00:32:46.210 and how that would impact
NOTE Confidence: 0.9770594

00:32:46.210 --> 00:32:47.890 oral care. So we wanted
NOTE Confidence: 0.9770594

00:32:47.890 --> 00:32:48.610 to see is that is
NOTE Confidence: 0.9770594

00:32:48.610 --> 00:32:49.490 it possible for us to
NOTE Confidence: 0.9770594

00:32:49.490 --> 00:32:51.465 use machine learning to accelerate

NOTE Confidence: 0.9770594
00:32:51.465 --> 00:32:52.585 this a little bit, at
NOTE Confidence: 0.9770594
00:32:52.585 --> 00:32:54.684 least find regions within the
NOTE Confidence: 0.9770594
00:32:54.905 --> 00:32:55.625 within the slide,
NOTE Confidence: 0.9456598
00:32:56.585 --> 00:32:57.085 that,
NOTE Confidence: 0.9367885
00:32:58.665 --> 00:32:59.705 that that a pathologist can
NOTE Confidence: 0.9367885
00:32:59.705 --> 00:33:00.985 then look at. So we
NOTE Confidence: 0.9367885
00:33:00.985 --> 00:33:02.105 studied this based on two
NOTE Confidence: 0.9367885
00:33:02.105 --> 00:33:03.304 cohorts. One of them was
NOTE Confidence: 0.9367885
00:33:03.304 --> 00:33:04.365 collected at Harvard,
NOTE Confidence: 0.89331084
00:33:05.190 --> 00:33:06.710 and it it was scanned
NOTE Confidence: 0.89331084
00:33:06.710 --> 00:33:08.169 using a micro CT scanner.
NOTE Confidence: 0.93846047
00:33:09.909 --> 00:33:10.950 And that's this cohort. And
NOTE Confidence: 0.93846047
00:33:10.950 --> 00:33:12.230 then there's another cohort that
NOTE Confidence: 0.93846047
00:33:12.230 --> 00:33:13.690 came from Jonathan Liu's group,
NOTE Confidence: 0.93846047
00:33:13.909 --> 00:33:14.950 at at the University of
NOTE Confidence: 0.93846047

00:33:14.950 --> 00:33:16.710 Washington. But, basically, we developed
NOTE Confidence: 0.93846047
00:33:16.710 --> 00:33:17.110 a,
NOTE Confidence: 0.8311669
00:33:17.510 --> 00:33:18.890 a multi instance learning
NOTE Confidence: 0.93555903
00:33:19.304 --> 00:33:20.664 based framework that was adopted
NOTE Confidence: 0.93555903
00:33:20.664 --> 00:33:21.625 for three d three d
NOTE Confidence: 0.93555903
00:33:21.625 --> 00:33:23.465 pathology. So much more compute
NOTE Confidence: 0.93555903
00:33:23.465 --> 00:33:23.965 intensive,
NOTE Confidence: 0.9592579
00:33:24.664 --> 00:33:25.945 using three d patches or
NOTE Confidence: 0.9592579
00:33:25.945 --> 00:33:27.384 voxels instead of instead of
NOTE Confidence: 0.9592579
00:33:27.384 --> 00:33:29.144 two dimensional patches, feature extraction
NOTE Confidence: 0.9592579
00:33:29.144 --> 00:33:30.105 in in three d and
NOTE Confidence: 0.9592579
00:33:30.105 --> 00:33:31.705 eventually eventually feature aggregation in
NOTE Confidence: 0.9592579
00:33:31.705 --> 00:33:32.710 three d. And we did
NOTE Confidence: 0.9592579
00:33:32.710 --> 00:33:33.670 quite a lot of analysis
NOTE Confidence: 0.9592579
00:33:33.670 --> 00:33:34.790 around this, figuring out what
NOTE Confidence: 0.9592579
00:33:34.790 --> 00:33:35.990 the best setup would be

NOTE Confidence: 0.9592579
00:33:35.990 --> 00:33:36.950 when this large amount of
NOTE Confidence: 0.9592579
00:33:36.950 --> 00:33:37.750 data would be would be
NOTE Confidence: 0.9592579
00:33:37.750 --> 00:33:39.190 available. This this article was
NOTE Confidence: 0.9592579
00:33:39.190 --> 00:33:40.550 published just a couple of
NOTE Confidence: 0.9592579
00:33:40.550 --> 00:33:41.210 months ago.
NOTE Confidence: 0.9591901
00:33:41.750 --> 00:33:42.870 And we showed that as
NOTE Confidence: 0.9591901
00:33:42.870 --> 00:33:43.930 you use increased,
NOTE Confidence: 0.9994393
00:33:45.110 --> 00:33:45.610 volume
NOTE Confidence: 0.9584649
00:33:46.365 --> 00:33:47.885 of tissue, the model did
NOTE Confidence: 0.9584649
00:33:47.885 --> 00:33:48.385 did
NOTE Confidence: 0.9409044
00:33:48.765 --> 00:33:49.725 did better in terms of
NOTE Confidence: 0.9409044
00:33:49.725 --> 00:33:51.805 separating patients into into distinct
NOTE Confidence: 0.9409044
00:33:51.805 --> 00:33:53.265 risk risk groups.
NOTE Confidence: 0.9922416
00:33:54.605 --> 00:33:55.645 And, of course, you can
NOTE Confidence: 0.9922416
00:33:55.645 --> 00:33:56.605 then go in and look
NOTE Confidence: 0.9922416

00:33:56.605 --> 00:33:58.285 at what's most important within

NOTE Confidence: 0.9922416

00:33:58.285 --> 00:33:59.425 the within the entire

NOTE Confidence: 0.99969363

00:33:59.820 --> 00:34:00.320 volume

NOTE Confidence: 0.9531636

00:34:00.779 --> 00:34:02.460 of this, or or or

NOTE Confidence: 0.9531636

00:34:02.460 --> 00:34:03.580 this entire three d three

NOTE Confidence: 0.9531636

00:34:03.580 --> 00:34:04.940 d volume. We're continuing to

NOTE Confidence: 0.9531636

00:34:04.940 --> 00:34:06.700 investigate this for other other

NOTE Confidence: 0.9531636

00:34:06.700 --> 00:34:08.319 diseases, and we have a,

NOTE Confidence: 0.7938933

00:34:10.780 --> 00:34:11.920 a recent large,

NOTE Confidence: 0.9260875

00:34:12.460 --> 00:34:13.440 national effort,

NOTE Confidence: 0.9991835

00:34:14.295 --> 00:34:15.835 to use this for precision

NOTE Confidence: 0.9991835

00:34:15.895 --> 00:34:16.954 surgical interventions.

NOTE Confidence: 0.95912236

00:34:22.135 --> 00:34:23.415 So the next thing I'll

NOTE Confidence: 0.95912236

00:34:23.415 --> 00:34:24.775 quickly touch upon is some

NOTE Confidence: 0.95912236

00:34:24.775 --> 00:34:25.494 of the new work we

NOTE Confidence: 0.95912236

00:34:25.494 --> 00:34:26.454 are doing in trying to

NOTE Confidence: 0.95912236
00:34:26.454 --> 00:34:27.839 do AI driven three d
NOTE Confidence: 0.95912236
00:34:27.839 --> 00:34:29.440 spatial transcriptomics. Now a question
NOTE Confidence: 0.95912236
00:34:29.440 --> 00:34:30.880 mark there because this is
NOTE Confidence: 0.95912236
00:34:30.880 --> 00:34:32.719 relatively very new work, and
NOTE Confidence: 0.95912236
00:34:32.719 --> 00:34:34.180 we're still not quite sure
NOTE Confidence: 0.95912236
00:34:34.239 --> 00:34:35.200 what we're gonna find. But
NOTE Confidence: 0.95912236
00:34:35.200 --> 00:34:36.079 I wanted to share some
NOTE Confidence: 0.95912236
00:34:36.079 --> 00:34:37.779 of these early results,
NOTE Confidence: 0.99967855
00:34:38.160 --> 00:34:38.900 here today.
NOTE Confidence: 0.9803601
00:34:39.759 --> 00:34:40.719 So with the with the
NOTE Confidence: 0.9803601
00:34:40.719 --> 00:34:42.000 micro CT scanning or with
NOTE Confidence: 0.9803601
00:34:42.000 --> 00:34:42.855 the open top light sheet
NOTE Confidence: 0.9803601
00:34:42.935 --> 00:34:45.094 microscopy scanning, we have very
NOTE Confidence: 0.9803601
00:34:45.094 --> 00:34:46.635 nice volumes of
NOTE Confidence: 0.9874076
00:34:46.935 --> 00:34:47.655 of of of tissue that
NOTE Confidence: 0.9874076

00:34:47.655 --> 00:34:48.375 we can we can look
NOTE Confidence: 0.9874076

00:34:48.375 --> 00:34:49.494 at. And there have been
NOTE Confidence: 0.9874076

00:34:49.494 --> 00:34:51.015 a multitude of studies that
NOTE Confidence: 0.9874076

00:34:51.015 --> 00:34:52.055 have shown that you can
NOTE Confidence: 0.9874076

00:34:52.055 --> 00:34:52.555 predict
NOTE Confidence: 0.95773613

00:34:53.015 --> 00:34:53.515 spatial,
NOTE Confidence: 0.98554885

00:34:54.695 --> 00:34:55.515 spatial transcriptomics
NOTE Confidence: 0.9412388

00:34:55.975 --> 00:34:57.739 from histology alone at least
NOTE Confidence: 0.9412388

00:34:57.739 --> 00:34:58.540 to at least to a
NOTE Confidence: 0.9412388

00:34:58.540 --> 00:34:59.739 degree. So we wanted to
NOTE Confidence: 0.9412388

00:34:59.739 --> 00:35:00.859 see if we can leverage
NOTE Confidence: 0.9412388

00:35:00.859 --> 00:35:02.300 that to to predict ST
NOTE Confidence: 0.9412388

00:35:02.300 --> 00:35:03.119 in three d.
NOTE Confidence: 0.9455375

00:35:03.980 --> 00:35:04.480 But,
NOTE Confidence: 0.96632665

00:35:07.020 --> 00:35:08.060 of course, we can build
NOTE Confidence: 0.96632665

00:35:08.060 --> 00:35:09.099 these models based on lots

NOTE Confidence: 0.96632665
00:35:09.099 --> 00:35:10.219 of historical data, but we
NOTE Confidence: 0.96632665
00:35:10.219 --> 00:35:11.260 all also also wanted to
NOTE Confidence: 0.96632665
00:35:11.260 --> 00:35:12.219 see if it's possible for
NOTE Confidence: 0.96632665
00:35:12.219 --> 00:35:13.635 us to do some inpatient
NOTE Confidence: 0.96632665
00:35:13.635 --> 00:35:15.495 fine tuning. So a new
NOTE Confidence: 0.96632665
00:35:15.635 --> 00:35:16.755 block, we image it with
NOTE Confidence: 0.96632665
00:35:16.755 --> 00:35:18.114 CT. We do some do
NOTE Confidence: 0.96632665
00:35:18.114 --> 00:35:19.555 do some spatial transcriptomics in
NOTE Confidence: 0.96632665
00:35:19.555 --> 00:35:20.675 the top and bottom and
NOTE Confidence: 0.96632665
00:35:20.675 --> 00:35:21.955 fine tune this existing network
NOTE Confidence: 0.96632665
00:35:21.955 --> 00:35:23.155 that's probably been trained on
NOTE Confidence: 0.96632665
00:35:23.155 --> 00:35:24.594 lots and lots of data
NOTE Confidence: 0.96632665
00:35:24.594 --> 00:35:25.895 from the same tissue
NOTE Confidence: 0.94305
00:35:26.195 --> 00:35:27.395 and and and ST from
NOTE Confidence: 0.94305
00:35:27.395 --> 00:35:28.135 other patients.
NOTE Confidence: 0.9847162

00:35:28.489 --> 00:35:29.849 And this inpatient fine tuning
NOTE Confidence: 0.9847162

00:35:29.849 --> 00:35:31.130 could potentially then lead to
NOTE Confidence: 0.9847162

00:35:31.130 --> 00:35:32.010 lead to better,
NOTE Confidence: 0.94884133

00:35:32.650 --> 00:35:34.170 better interpolation in in three
NOTE Confidence: 0.94884133

00:35:34.170 --> 00:35:35.770 dimensions. It's just a hypothesis,
NOTE Confidence: 0.94884133

00:35:35.770 --> 00:35:36.650 and we wanted to see
NOTE Confidence: 0.94884133

00:35:36.650 --> 00:35:37.450 if we could we could
NOTE Confidence: 0.94884133

00:35:37.450 --> 00:35:38.730 test this. There are some
NOTE Confidence: 0.94884133

00:35:38.730 --> 00:35:40.089 some initial results. We use
NOTE Confidence: 0.94884133

00:35:40.089 --> 00:35:41.289 a number of different mechanisms
NOTE Confidence: 0.94884133

00:35:41.289 --> 00:35:42.089 to see what the best
NOTE Confidence: 0.94884133

00:35:42.089 --> 00:35:43.364 approach would be to to
NOTE Confidence: 0.94884133

00:35:43.364 --> 00:35:43.864 predict,
NOTE Confidence: 0.959101

00:35:44.245 --> 00:35:46.165 spatial transcriptomics from from from
NOTE Confidence: 0.959101

00:35:46.165 --> 00:35:46.985 h and e.
NOTE Confidence: 0.9945328

00:35:47.765 --> 00:35:48.745 And we tried

NOTE Confidence: 0.89258456
00:35:49.205 --> 00:35:50.085 a a number of different
NOTE Confidence: 0.89258456
00:35:50.085 --> 00:35:51.945 techniques, but contrasting between,
NOTE Confidence: 0.96513176
00:35:52.645 --> 00:35:53.525 ST and h and e
NOTE Confidence: 0.96513176
00:35:53.525 --> 00:35:54.965 seems to be, the best
NOTE Confidence: 0.96513176
00:35:54.965 --> 00:35:56.405 approach. We built a predictor.
NOTE Confidence: 0.96513176
00:35:56.405 --> 00:35:58.119 And then also incorporating some
NOTE Confidence: 0.96513176
00:35:58.119 --> 00:35:59.900 depth information leads to improved
NOTE Confidence: 0.96513176
00:36:00.040 --> 00:36:01.739 improved performance as well.
NOTE Confidence: 0.94970864
00:36:02.280 --> 00:36:03.800 But eventually, we're able to
NOTE Confidence: 0.94970864
00:36:03.800 --> 00:36:04.619 get to,
NOTE Confidence: 0.98116803
00:36:05.400 --> 00:36:06.540 a point where we can
NOTE Confidence: 0.98116803
00:36:06.760 --> 00:36:08.040 interpolate this in in three
NOTE Confidence: 0.98116803
00:36:08.040 --> 00:36:09.560 dimension and also confirm that
NOTE Confidence: 0.98116803
00:36:09.560 --> 00:36:10.680 that that this is correct
NOTE Confidence: 0.98116803
00:36:10.680 --> 00:36:12.135 to a degree. We've expanded
NOTE Confidence: 0.98116803

00:36:12.135 --> 00:36:13.415 this to three different disease
NOTE Confidence: 0.98116803

00:36:13.415 --> 00:36:14.614 models. This this example is
NOTE Confidence: 0.98116803

00:36:14.614 --> 00:36:15.435 on the prostate.
NOTE Confidence: 0.93520695

00:36:15.815 --> 00:36:17.655 We've expanded it just expanded
NOTE Confidence: 0.93520695

00:36:17.655 --> 00:36:18.775 this to three different disease
NOTE Confidence: 0.93520695

00:36:18.775 --> 00:36:19.975 models as well as a
NOTE Confidence: 0.93520695

00:36:19.975 --> 00:36:21.255 lot of additional data that
NOTE Confidence: 0.93520695

00:36:21.255 --> 00:36:21.915 was available
NOTE Confidence: 0.955907

00:36:22.295 --> 00:36:23.815 with our collaborators at the
NOTE Confidence: 0.955907

00:36:23.815 --> 00:36:24.955 Broad and other places,
NOTE Confidence: 0.99168044

00:36:26.390 --> 00:36:27.510 and are sort of continuing
NOTE Confidence: 0.99168044

00:36:27.510 --> 00:36:28.230 to work in this,
NOTE Confidence: 0.8605808

00:36:28.790 --> 00:36:30.010 in in in this direction.
NOTE Confidence: 0.9654672

00:36:31.190 --> 00:36:32.150 The last thing I wanna
NOTE Confidence: 0.9654672

00:36:32.150 --> 00:36:33.350 touch upon is is bias
NOTE Confidence: 0.9654672

00:36:33.350 --> 00:36:35.030 and fairness in computational pathology

NOTE Confidence: 0.9654672
00:36:35.030 --> 00:36:36.010 datasets. So,
NOTE Confidence: 0.9677306
00:36:37.030 --> 00:36:38.230 a lot of computational data,
NOTE Confidence: 0.9677306
00:36:38.390 --> 00:36:39.750 pathology data comes from large
NOTE Confidence: 0.9677306
00:36:39.750 --> 00:36:41.105 academic medical centers, and the
NOTE Confidence: 0.9677306
00:36:41.105 --> 00:36:42.305 data is not very diverse.
NOTE Confidence: 0.9677306
00:36:42.305 --> 00:36:43.585 And and we have done
NOTE Confidence: 0.9677306
00:36:43.585 --> 00:36:44.944 some analysis where we train
NOTE Confidence: 0.9677306
00:36:44.944 --> 00:36:46.065 on on on large cohorts
NOTE Confidence: 0.9677306
00:36:46.065 --> 00:36:47.105 of data that are commonly
NOTE Confidence: 0.9677306
00:36:47.105 --> 00:36:48.545 used in model development like
NOTE Confidence: 0.9677306
00:36:48.545 --> 00:36:49.744 the TCGA as well as
NOTE Confidence: 0.9677306
00:36:49.744 --> 00:36:51.265 internal data. And what happens
NOTE Confidence: 0.9677306
00:36:51.265 --> 00:36:52.484 when you adapt to,
NOTE Confidence: 0.98333305
00:36:53.200 --> 00:36:54.960 data that is independent and
NOTE Confidence: 0.98333305
00:36:54.960 --> 00:36:56.400 stratified by race or other
NOTE Confidence: 0.98333305

00:36:56.400 --> 00:36:57.300 protected subgroups.
NOTE Confidence: 0.93018776

00:36:58.080 --> 00:36:58.880 And that's what led to
NOTE Confidence: 0.93018776

00:36:58.880 --> 00:37:00.400 this study where we it
NOTE Confidence: 0.93018776

00:37:00.400 --> 00:37:01.600 was also published earlier this
NOTE Confidence: 0.93018776

00:37:01.600 --> 00:37:02.900 year where we wanted to
NOTE Confidence: 0.93018776

00:37:02.960 --> 00:37:05.140 investigate demographic shifts and misdiagnosis
NOTE Confidence: 0.93018776

00:37:05.360 --> 00:37:06.555 by computational pathology
NOTE Confidence: 0.9744052

00:37:06.954 --> 00:37:07.454 models.
NOTE Confidence: 0.9034781

00:37:08.234 --> 00:37:09.675 Overall, the idea is that
NOTE Confidence: 0.9034781

00:37:09.675 --> 00:37:10.234 is that,
NOTE Confidence: 0.97381127

00:37:10.714 --> 00:37:11.594 when when you train a
NOTE Confidence: 0.97381127

00:37:11.594 --> 00:37:12.795 model on a specific cohort
NOTE Confidence: 0.97381127

00:37:12.795 --> 00:37:13.835 when and you transfer it
NOTE Confidence: 0.97381127

00:37:13.835 --> 00:37:14.335 to,
NOTE Confidence: 0.94197863

00:37:14.714 --> 00:37:15.434 to to new kinds of
NOTE Confidence: 0.94197863

00:37:15.434 --> 00:37:16.714 data, it often does not

NOTE Confidence: 0.94197863
00:37:16.714 --> 00:37:17.835 adapt. And this issue around
NOTE Confidence: 0.94197863
00:37:17.835 --> 00:37:19.515 domain adaptation is is like
NOTE Confidence: 0.94197863
00:37:19.515 --> 00:37:20.859 a age old problem, but
NOTE Confidence: 0.94197863
00:37:20.859 --> 00:37:21.980 it's sort of exaggerated in
NOTE Confidence: 0.94197863
00:37:21.980 --> 00:37:22.140 health care.
NOTE Confidence: 0.9853319
00:37:23.500 --> 00:37:24.859 And there are large differences,
NOTE Confidence: 0.9853319
00:37:24.859 --> 00:37:26.480 for example, in in how
NOTE Confidence: 0.9853319
00:37:26.619 --> 00:37:28.640 the individual scanner behave.
NOTE Confidence: 0.9143975
00:37:29.500 --> 00:37:31.020 This particular slide was scanned
NOTE Confidence: 0.9143975
00:37:31.020 --> 00:37:32.380 on, like, a Haohomatsu scanner
NOTE Confidence: 0.9143975
00:37:32.380 --> 00:37:33.260 and a period scanner and
NOTE Confidence: 0.9143975
00:37:33.260 --> 00:37:34.765 three d stack scanner, leading
NOTE Confidence: 0.9143975
00:37:34.765 --> 00:37:35.645 to a very different color
NOTE Confidence: 0.9143975
00:37:35.645 --> 00:37:36.925 gamut. The same is true
NOTE Confidence: 0.9143975
00:37:36.925 --> 00:37:38.785 in radiology radiology as well.
NOTE Confidence: 0.9143975

00:37:38.844 --> 00:37:39.505 The datasets,
NOTE Confidence: 0.9759659

00:37:40.125 --> 00:37:40.945 also have,
NOTE Confidence: 0.965023

00:37:42.445 --> 00:37:43.405 you know, they're they're very
NOTE Confidence: 0.965023

00:37:43.405 --> 00:37:44.525 consistent. They don't have a
NOTE Confidence: 0.965023

00:37:44.525 --> 00:37:45.665 lot of lot of diversity.
NOTE Confidence: 0.7587725

00:37:46.045 --> 00:37:46.205 The,
NOTE Confidence: 0.96479297

00:37:47.180 --> 00:37:48.380 point we really wanted to
NOTE Confidence: 0.96479297

00:37:48.380 --> 00:37:49.900 investigate was that can we
NOTE Confidence: 0.96479297

00:37:49.900 --> 00:37:51.020 go in and look at
NOTE Confidence: 0.96479297

00:37:51.020 --> 00:37:53.100 each and every possible modeling
NOTE Confidence: 0.96479297

00:37:53.100 --> 00:37:54.320 choice that people make,
NOTE Confidence: 0.88460827

00:37:54.780 --> 00:37:57.200 when designing their computational pathology
NOTE Confidence: 0.88460827

00:37:57.260 --> 00:37:58.540 setup or or their or
NOTE Confidence: 0.88460827

00:37:58.540 --> 00:37:59.680 their training architecture,
NOTE Confidence: 0.97677344

00:38:00.925 --> 00:38:02.625 and see how that impacts,
NOTE Confidence: 0.9505309

00:38:04.045 --> 00:38:04.704 the outcome,

NOTE Confidence: 0.9577663

00:38:05.885 --> 00:38:07.244 and the the the the

NOTE Confidence: 0.9577663

00:38:07.244 --> 00:38:09.665 classification results when it's stratified

NOTE Confidence: 0.9577663

00:38:09.805 --> 00:38:11.425 by some of these protected

NOTE Confidence: 0.9577663

00:38:11.484 --> 00:38:13.744 protected subgroups. So we varied

NOTE Confidence: 0.9577663

00:38:13.805 --> 00:38:15.410 some all the different preprocessing

NOTE Confidence: 0.9577663

00:38:15.630 --> 00:38:17.569 techniques, the utility of

NOTE Confidence: 0.99567753

00:38:17.869 --> 00:38:18.369 the

NOTE Confidence: 0.9624534

00:38:18.910 --> 00:38:20.270 the the foundation model or

NOTE Confidence: 0.9624534

00:38:20.270 --> 00:38:21.630 the self supervised model that

NOTE Confidence: 0.9624534

00:38:21.630 --> 00:38:22.829 was that was used, as

NOTE Confidence: 0.9624534

00:38:22.829 --> 00:38:24.190 well as some tricks that

NOTE Confidence: 0.9624534

00:38:24.190 --> 00:38:25.969 are commonly used to,

NOTE Confidence: 0.94147986

00:38:26.430 --> 00:38:28.450 improve fairness, like adversarial regularization

NOTE Confidence: 0.91487247

00:38:28.750 --> 00:38:29.650 and so forth,

NOTE Confidence: 0.923153

00:38:30.275 --> 00:38:31.875 and other fairness fairness strategies.

NOTE Confidence: 0.923153

00:38:31.875 --> 00:38:33.235 And overall, we found that
NOTE Confidence: 0.923153

00:38:33.235 --> 00:38:34.675 the most important component is
NOTE Confidence: 0.923153

00:38:34.675 --> 00:38:35.795 that how rich your feature
NOTE Confidence: 0.923153

00:38:35.955 --> 00:38:36.855 features are.
NOTE Confidence: 0.9594325

00:38:37.315 --> 00:38:38.995 And by using, only features
NOTE Confidence: 0.9594325

00:38:38.995 --> 00:38:39.795 in this case, we're able
NOTE Confidence: 0.9594325

00:38:39.795 --> 00:38:41.015 to improve performance
NOTE Confidence: 0.9858931

00:38:41.475 --> 00:38:42.855 for a number of these,
NOTE Confidence: 0.9047891

00:38:43.155 --> 00:38:43.655 different,
NOTE Confidence: 0.94903016

00:38:44.995 --> 00:38:46.870 different classification examples that were
NOTE Confidence: 0.94903016

00:38:46.870 --> 00:38:47.590 that were used in the
NOTE Confidence: 0.94903016

00:38:47.590 --> 00:38:49.350 in the study. So I'll,
NOTE Confidence: 0.97833145

00:38:49.830 --> 00:38:51.110 stop here in the interest
NOTE Confidence: 0.97833145

00:38:51.110 --> 00:38:52.250 of time, but I do
NOTE Confidence: 0.9658065

00:38:52.550 --> 00:38:53.989 wanna read from this poem
NOTE Confidence: 0.9658065

00:38:53.989 --> 00:38:55.670 that Judith Prevett wrote. She

NOTE Confidence: 0.9658065

00:38:55.670 --> 00:38:56.730 was one of the

NOTE Confidence: 0.95248926

00:38:57.525 --> 00:38:59.684 pioneers in analyzing microscopy images

NOTE Confidence: 0.95248926

00:38:59.684 --> 00:39:01.125 in in computers, some some

NOTE Confidence: 0.95248926

00:39:01.125 --> 00:39:02.244 really pioneering work in the

NOTE Confidence: 0.95248926

00:39:02.244 --> 00:39:04.325 1960s and 70s. She writes

NOTE Confidence: 0.95248926

00:39:04.325 --> 00:39:06.085 that optical illusions can deceive

NOTE Confidence: 0.95248926

00:39:06.085 --> 00:39:07.605 the subjective eye, but objective

NOTE Confidence: 0.95248926

00:39:07.605 --> 00:39:08.664 measurements and algorithms

NOTE Confidence: 0.9982408

00:39:09.045 --> 00:39:10.505 are assumed not to lie.

NOTE Confidence: 0.9982408

00:39:10.530 --> 00:39:11.810 It's often said that medicine

NOTE Confidence: 0.9982408

00:39:11.810 --> 00:39:13.410 could use such objectivity and

NOTE Confidence: 0.9982408

00:39:13.410 --> 00:39:14.550 thought that this justifies

NOTE Confidence: 0.98673093

00:39:14.850 --> 00:39:17.330 machine intelligence activity. Artificial intelligence

NOTE Confidence: 0.98673093

00:39:17.330 --> 00:39:18.930 is another craze that uses

NOTE Confidence: 0.98673093

00:39:18.930 --> 00:39:19.969 computers to cope with the

NOTE Confidence: 0.98673093

00:39:19.969 --> 00:39:21.969 diagnostic maze. Though the criteria

NOTE Confidence: 0.98673093

00:39:21.969 --> 00:39:23.250 for intelligence has never been

NOTE Confidence: 0.98673093

00:39:23.250 --> 00:39:25.195 resolved, paper after paper claims,

NOTE Confidence: 0.98673093

00:39:25.195 --> 00:39:26.315 the problem has already been

NOTE Confidence: 0.98673093

00:39:26.315 --> 00:39:27.594 solved. So we still have

NOTE Confidence: 0.98673093

00:39:27.594 --> 00:39:29.135 a way to go before

NOTE Confidence: 0.97500706

00:39:29.594 --> 00:39:30.714 we we address some of

NOTE Confidence: 0.97500706

00:39:30.714 --> 00:39:32.175 these, critical issues,

NOTE Confidence: 0.97259563

00:39:33.515 --> 00:39:34.655 that that that exist,

NOTE Confidence: 0.97596765

00:39:35.195 --> 00:39:36.015 with developing

NOTE Confidence: 0.9819446

00:39:36.489 --> 00:39:38.330 effective machine learning algorithms for

NOTE Confidence: 0.9819446

00:39:38.330 --> 00:39:39.530 for pathology. And I'd like

NOTE Confidence: 0.9819446

00:39:39.530 --> 00:39:40.810 to thank all the funding

NOTE Confidence: 0.9819446

00:39:40.810 --> 00:39:41.770 that we received to do

NOTE Confidence: 0.9819446

00:39:41.770 --> 00:39:42.510 this work,

NOTE Confidence: 0.9183416

00:39:43.130 --> 00:39:43.930 as well as all the
NOTE Confidence: 0.9183416

00:39:43.930 --> 00:39:45.610 PhD students, postdocs who have
NOTE Confidence: 0.9183416

00:39:45.610 --> 00:39:46.910 worked worked in the lab.
NOTE Confidence: 0.74376875

00:39:56.305 --> 00:39:58.385 Yes. Another question. So we
NOTE Confidence: 0.74376875

00:39:58.385 --> 00:39:59.505 know that PR can have
NOTE Confidence: 0.74376875

00:39:59.505 --> 00:40:00.005 hallucination.
NOTE Confidence: 0.72874385

00:40:00.625 --> 00:40:01.665 Yeah. Always come up with
NOTE Confidence: 0.72874385

00:40:01.665 --> 00:40:03.425 Hunter. What if the same
NOTE Confidence: 0.72874385

00:40:03.425 --> 00:40:05.045 car make that Yeah. Policy
NOTE Confidence: 0.72874385

00:40:05.185 --> 00:40:06.140 share Yeah.
NOTE Confidence: 0.92413986

00:40:15.640 --> 00:40:16.920 Yeah. Well, we are trying
NOTE Confidence: 0.92413986

00:40:16.920 --> 00:40:17.655 to get it to stop.
NOTE Confidence: 0.92413986

00:40:17.735 --> 00:40:19.575 So so the multimodal large
NOTE Confidence: 0.92413986

00:40:19.575 --> 00:40:20.935 language model that we built
NOTE Confidence: 0.92413986

00:40:20.935 --> 00:40:21.995 that is that is prone
NOTE Confidence: 0.92413986

00:40:22.295 --> 00:40:22.875 to hallucinations,

NOTE Confidence: 0.9700358
00:40:23.655 --> 00:40:24.614 we built in a lot
NOTE Confidence: 0.9700358
00:40:24.614 --> 00:40:25.975 of guardrails. So it would
NOTE Confidence: 0.9700358
00:40:25.975 --> 00:40:27.035 stop giving a,
NOTE Confidence: 0.9796047
00:40:28.455 --> 00:40:29.735 stop from making a diagnosis
NOTE Confidence: 0.9796047
00:40:29.735 --> 00:40:30.795 if it's not sure.
NOTE Confidence: 0.9554329
00:40:31.910 --> 00:40:32.950 Or, you know, if you
NOTE Confidence: 0.9554329
00:40:32.950 --> 00:40:34.150 give it an image, a
NOTE Confidence: 0.9554329
00:40:34.150 --> 00:40:35.350 model that's entirely trained on
NOTE Confidence: 0.9554329
00:40:35.350 --> 00:40:36.310 pathology images and you give
NOTE Confidence: 0.9554329
00:40:36.310 --> 00:40:37.030 it an image of a
NOTE Confidence: 0.9554329
00:40:37.030 --> 00:40:38.230 cat, it would still say
NOTE Confidence: 0.9554329
00:40:38.230 --> 00:40:39.610 maybe it's squamous cell carcinoma.
NOTE Confidence: 0.93528473
00:40:40.870 --> 00:40:42.230 So so you don't want
NOTE Confidence: 0.93528473
00:40:42.230 --> 00:40:43.450 it to do that. So
NOTE Confidence: 0.97200906
00:40:43.935 --> 00:40:45.215 we build in guardrails that
NOTE Confidence: 0.97200906

00:40:45.215 --> 00:40:46.275 that prevents hallucinations.
NOTE Confidence: 0.9994508

00:40:46.895 --> 00:40:48.114 More data for training
NOTE Confidence: 0.8913075

00:40:48.575 --> 00:40:49.075 prevents,
NOTE Confidence: 0.8534353

00:40:49.614 --> 00:40:50.114 hallucinations.
NOTE Confidence: 0.9738948

00:40:50.975 --> 00:40:53.135 Better pretraining data also prevents
NOTE Confidence: 0.9738948

00:40:53.135 --> 00:40:53.635 hallucinations.
NOTE Confidence: 0.50365514

00:40:54.455 --> 00:40:54.955 I
NOTE Confidence: 0.21440892

00:40:55.775 --> 00:40:57.075 I actually seen that.
NOTE Confidence: 0.624324

00:41:14.364 --> 00:41:15.965 Never received the answers they
NOTE Confidence: 0.624324

00:41:15.965 --> 00:41:17.885 have. Yeah. Always come with
NOTE Confidence: 0.624324

00:41:17.885 --> 00:41:18.385 the.
NOTE Confidence: 0.7461332

00:41:18.844 --> 00:41:20.605 Right? Yeah. Back to the
NOTE Confidence: 0.7461332

00:41:20.605 --> 00:41:22.465 most question, but Yep. Okay.
NOTE Confidence: 0.92439514

00:41:23.005 --> 00:41:23.505 Safeguard
NOTE Confidence: 0.8071453

00:41:23.805 --> 00:41:25.344 to say that, you know,
NOTE Confidence: 0.7635096

00:41:25.805 --> 00:41:27.325 particularly with your with your

NOTE Confidence: 0.7635096
00:41:27.325 --> 00:41:28.510 past chat. Yeah. Do you
NOTE Confidence: 0.7635096
00:41:28.510 --> 00:41:29.650 have, like, seen forever,
NOTE Confidence: 0.7847828
00:41:30.190 --> 00:41:31.470 pets check and say, I
NOTE Confidence: 0.7847828
00:41:31.470 --> 00:41:32.370 don't know that?
NOTE Confidence: 0.9378186
00:41:33.630 --> 00:41:34.670 We are trying to build
NOTE Confidence: 0.9378186
00:41:34.670 --> 00:41:36.030 in those guardrails. Right? So
NOTE Confidence: 0.9378186
00:41:36.030 --> 00:41:37.150 so the the there's this
NOTE Confidence: 0.9378186
00:41:37.150 --> 00:41:38.989 whole field of, study in
NOTE Confidence: 0.9378186
00:41:38.989 --> 00:41:40.685 machine learning where you get
NOTE Confidence: 0.9378186
00:41:40.685 --> 00:41:41.965 the model to abstain from
NOTE Confidence: 0.9378186
00:41:41.965 --> 00:41:43.885 making predictions. Right? It's it's
NOTE Confidence: 0.9378186
00:41:43.885 --> 00:41:45.165 ongoing research in that area
NOTE Confidence: 0.9378186
00:41:45.165 --> 00:41:46.125 that how do you abstain
NOTE Confidence: 0.9378186
00:41:46.125 --> 00:41:47.405 from making a prediction. If
NOTE Confidence: 0.9378186
00:41:47.405 --> 00:41:48.445 you're not sure, how do
NOTE Confidence: 0.9378186

00:41:48.525 --> 00:41:49.665 how does the model abstain?
NOTE Confidence: 0.68142253

00:41:49.965 --> 00:41:51.485 So the the the latest
NOTE Confidence: 0.68142253

00:41:51.485 --> 00:41:53.425 question for the patch check.
NOTE Confidence: 0.68142253

00:41:53.485 --> 00:41:54.525 So how wide are they
NOTE Confidence: 0.68142253

00:41:54.525 --> 00:41:55.025 available
NOTE Confidence: 0.90333897

00:41:56.130 --> 00:41:57.489 Your patch check within your
NOTE Confidence: 0.90333897

00:41:57.489 --> 00:41:59.170 department and Google. So what's
NOTE Confidence: 0.90333897

00:41:59.170 --> 00:41:59.969 the kind of Yeah. We
NOTE Confidence: 0.90333897

00:41:59.969 --> 00:42:00.849 we we have about a
NOTE Confidence: 0.90333897

00:42:00.849 --> 00:42:02.690 hundred people using it. The
NOTE Confidence: 0.90333897

00:42:02.690 --> 00:42:04.049 issue with making it available
NOTE Confidence: 0.90333897

00:42:04.049 --> 00:42:04.850 too widely is that it
NOTE Confidence: 0.90333897

00:42:04.930 --> 00:42:06.450 it's expensive. The deployment is
NOTE Confidence: 0.90333897

00:42:06.450 --> 00:42:07.969 is expensive because it it
NOTE Confidence: 0.90333897

00:42:07.969 --> 00:42:09.270 actively uses GPUs,
NOTE Confidence: 0.9813075

00:42:10.565 --> 00:42:11.204 in the,

NOTE Confidence: 0.9881513
00:42:11.525 --> 00:42:12.424 in the background.
NOTE Confidence: 0.94159055
00:42:13.765 --> 00:42:14.244 But,
NOTE Confidence: 0.9581751
00:42:14.565 --> 00:42:15.684 our group has spun off
NOTE Confidence: 0.9581751
00:42:15.684 --> 00:42:16.964 a startup company that that
NOTE Confidence: 0.9581751
00:42:16.964 --> 00:42:17.844 that plans to make it
NOTE Confidence: 0.9581751
00:42:17.844 --> 00:42:18.984 more widely available.
NOTE Confidence: 0.87862813
00:42:20.484 --> 00:42:21.305 And they're
NOTE Confidence: 0.9172026
00:42:21.940 --> 00:42:23.700 are they're expanding on the
NOTE Confidence: 0.9172026
00:42:23.700 --> 00:42:24.440 amount of data
NOTE Confidence: 0.9779164
00:42:25.140 --> 00:42:26.340 that they're using for training
NOTE Confidence: 0.9779164
00:42:26.340 --> 00:42:27.320 as well as evaluation.
NOTE Confidence: 0.93169874
00:42:28.340 --> 00:42:29.540 So hopefully, over time, I
NOTE Confidence: 0.93169874
00:42:29.540 --> 00:42:30.260 think it will it will
NOTE Confidence: 0.93169874
00:42:30.260 --> 00:42:31.000 become available.
NOTE Confidence: 0.96850467
00:43:03.325 --> 00:43:04.364 Yeah. Yeah. That's a that's
NOTE Confidence: 0.96850467

00:43:04.364 --> 00:43:05.485 a that's a great great
NOTE Confidence: 0.96850467

00:43:05.485 --> 00:43:06.685 question. There there are techniques
NOTE Confidence: 0.96850467

00:43:06.685 --> 00:43:07.885 you can use to to
NOTE Confidence: 0.96850467

00:43:07.885 --> 00:43:09.905 minimize the the batch effect.
NOTE Confidence: 0.8809506

00:43:11.005 --> 00:43:12.685 The spatial transcriptomic results that
NOTE Confidence: 0.8809506

00:43:12.685 --> 00:43:13.745 I showed,
NOTE Confidence: 0.98888457

00:43:14.525 --> 00:43:15.025 it's
NOTE Confidence: 0.9479195

00:43:15.440 --> 00:43:16.640 it's it's consistent. Like, so
NOTE Confidence: 0.9479195

00:43:16.640 --> 00:43:17.760 it's it's from within the
NOTE Confidence: 0.9479195

00:43:17.760 --> 00:43:18.260 same,
NOTE Confidence: 0.95081085

00:43:18.880 --> 00:43:20.560 data collection pipeline to to
NOTE Confidence: 0.95081085

00:43:20.560 --> 00:43:21.600 build a three d three
NOTE Confidence: 0.95081085

00:43:21.600 --> 00:43:22.260 d model.
NOTE Confidence: 0.9756561

00:43:22.720 --> 00:43:24.480 But, there's some other work
NOTE Confidence: 0.9756561

00:43:24.480 --> 00:43:25.780 from my group,
NOTE Confidence: 0.97588956

00:43:27.200 --> 00:43:29.120 the HEST benchmark and the

NOTE Confidence: 0.97588956
00:43:29.120 --> 00:43:31.075 corresponding library. It has some
NOTE Confidence: 0.97588956
00:43:31.075 --> 00:43:32.114 tools that you can use
NOTE Confidence: 0.97588956
00:43:32.114 --> 00:43:32.614 to
NOTE Confidence: 0.89854
00:43:32.995 --> 00:43:33.955 to to to reduce the
NOTE Confidence: 0.89854
00:43:33.955 --> 00:43:34.695 batch effect.
NOTE Confidence: 0.9175121
00:43:36.195 --> 00:43:37.475 That said, you probably cannot
NOTE Confidence: 0.9175121
00:43:37.475 --> 00:43:39.235 eliminate the the batch effect
NOTE Confidence: 0.9175121
00:43:39.235 --> 00:43:40.935 completely. Right? So you can
NOTE Confidence: 0.9175121
00:43:41.155 --> 00:43:42.535 probably reduce the batch
NOTE Confidence: 0.9589082
00:43:43.119 --> 00:43:44.660 effect that exists in,
NOTE Confidence: 0.9986688
00:43:45.520 --> 00:43:46.099 the image
NOTE Confidence: 0.9093359
00:43:46.560 --> 00:43:47.300 to a degree,
NOTE Confidence: 0.96455806
00:43:48.080 --> 00:43:49.200 to a degree you can
NOTE Confidence: 0.96455806
00:43:49.359 --> 00:43:50.800 if if it's it's something
NOTE Confidence: 0.96455806
00:43:50.800 --> 00:43:52.239 that's consistent across all of
NOTE Confidence: 0.96455806

00:43:52.239 --> 00:43:53.520 your data, you could you
NOTE Confidence: 0.96455806

00:43:53.520 --> 00:43:55.614 could perhaps eliminate that. But
NOTE Confidence: 0.96455806

00:43:55.614 --> 00:43:56.895 site specific batch effect, if
NOTE Confidence: 0.96455806

00:43:56.895 --> 00:43:57.695 you have lots of spatial
NOTE Confidence: 0.96455806

00:43:57.695 --> 00:43:59.315 transcriptomic data that you're bundling
NOTE Confidence: 0.96455806

00:43:59.375 --> 00:44:01.055 bundling together to perhaps train
NOTE Confidence: 0.96455806

00:44:01.055 --> 00:44:02.975 a contrastive model, very difficult
NOTE Confidence: 0.96455806

00:44:02.975 --> 00:44:03.555 to do.
NOTE Confidence: 0.84731025

00:45:02.655 --> 00:45:03.155 Yeah.
NOTE Confidence: 0.9234532

00:45:04.880 --> 00:45:06.320 Yeah. So the so your
NOTE Confidence: 0.9234532

00:45:06.320 --> 00:45:07.760 first question about the patch
NOTE Confidence: 0.9234532

00:45:07.760 --> 00:45:08.900 size. Right? So,
NOTE Confidence: 0.9964901

00:45:10.400 --> 00:45:11.360 there are a number of
NOTE Confidence: 0.9964901

00:45:11.360 --> 00:45:12.260 different ways,
NOTE Confidence: 0.9693181

00:45:12.640 --> 00:45:13.800 to think about this. The
NOTE Confidence: 0.9693181

00:45:13.800 --> 00:45:14.960 the the first thing is

NOTE Confidence: 0.9693181
00:45:14.960 --> 00:45:16.500 that the majority of computational
NOTE Confidence: 0.9693181
00:45:16.560 --> 00:45:18.234 pathology studies, they work at
NOTE Confidence: 0.9693181
00:45:18.234 --> 00:45:19.594 a single resolution. They patch
NOTE Confidence: 0.9693181
00:45:19.594 --> 00:45:20.555 everything out at two five
NOTE Confidence: 0.9693181
00:45:20.555 --> 00:45:21.775 six by two five six
NOTE Confidence: 0.9166157
00:45:22.075 --> 00:45:23.675 images. And in the in
NOTE Confidence: 0.9166157
00:45:23.675 --> 00:45:25.275 in the majority of cases,
NOTE Confidence: 0.9166157
00:45:25.275 --> 00:45:26.734 they work work just fine
NOTE Confidence: 0.8222858
00:45:27.114 --> 00:45:28.155 for whole slide of a
NOTE Confidence: 0.8222858
00:45:28.155 --> 00:45:29.455 classification. And that's because
NOTE Confidence: 0.9661645
00:45:30.119 --> 00:45:31.559 the morphologic feature that you're
NOTE Confidence: 0.9661645
00:45:31.559 --> 00:45:32.680 trying to identify is very
NOTE Confidence: 0.9661645
00:45:32.680 --> 00:45:33.719 clearly evident at a two
NOTE Confidence: 0.9661645
00:45:33.719 --> 00:45:34.599 five six by two five
NOTE Confidence: 0.9661645
00:45:34.599 --> 00:45:36.279 six patch. I've had pathologists
NOTE Confidence: 0.9661645

00:45:36.279 --> 00:45:37.000 tell me that, oh, I
NOTE Confidence: 0.9661645

00:45:37.000 --> 00:45:38.200 I can't identify something at
NOTE Confidence: 0.9661645

00:45:38.200 --> 00:45:38.920 a two five six by
NOTE Confidence: 0.9661645

00:45:38.920 --> 00:45:40.519 two five six patch. Why
NOTE Confidence: 0.9661645

00:45:40.519 --> 00:45:41.559 is this working so well
NOTE Confidence: 0.9661645

00:45:41.559 --> 00:45:42.765 even though each patch is
NOTE Confidence: 0.9661645

00:45:42.844 --> 00:45:43.744 is mutually exclusive
NOTE Confidence: 0.91862595

00:45:44.125 --> 00:45:45.644 and the model has no
NOTE Confidence: 0.91862595

00:45:45.644 --> 00:45:47.085 has no context? And that's
NOTE Confidence: 0.91862595

00:45:47.085 --> 00:45:48.844 perhaps because features are being,
NOTE Confidence: 0.91862595

00:45:49.244 --> 00:45:50.525 being extracted and then they're
NOTE Confidence: 0.91862595

00:45:50.525 --> 00:45:51.724 being aggregated so that that
NOTE Confidence: 0.91862595

00:45:51.724 --> 00:45:52.224 aggregation
NOTE Confidence: 0.91101277

00:45:53.404 --> 00:45:54.944 somehow counts for that.
NOTE Confidence: 0.96828246

00:45:55.404 --> 00:45:57.325 That said, intuitively, it doesn't
NOTE Confidence: 0.96828246

00:45:57.325 --> 00:45:58.305 make sense because

NOTE Confidence: 0.9997404
00:45:58.680 --> 00:45:59.180 the
NOTE Confidence: 0.99488246
00:45:59.640 --> 00:46:01.239 the the each patch is
NOTE Confidence: 0.99488246
00:46:01.239 --> 00:46:02.839 still mutually exclusive, and they're
NOTE Confidence: 0.99488246
00:46:02.839 --> 00:46:04.440 they're not linked together. There
NOTE Confidence: 0.99488246
00:46:04.440 --> 00:46:05.400 have been a multitude of
NOTE Confidence: 0.99488246
00:46:05.400 --> 00:46:06.760 studies using graphs and other
NOTE Confidence: 0.99488246
00:46:06.760 --> 00:46:08.040 techniques to link the patches
NOTE Confidence: 0.99488246
00:46:08.040 --> 00:46:09.640 together to improve context, but
NOTE Confidence: 0.99488246
00:46:09.640 --> 00:46:10.760 they haven't shown a substantial
NOTE Confidence: 0.99488246
00:46:10.760 --> 00:46:11.260 improvement
NOTE Confidence: 0.9701824
00:46:11.775 --> 00:46:13.234 in, in performance.
NOTE Confidence: 0.97644633
00:46:13.614 --> 00:46:14.815 That's that's one aspect. The
NOTE Confidence: 0.97644633
00:46:14.815 --> 00:46:15.855 other aspect is that the
NOTE Confidence: 0.97644633
00:46:15.855 --> 00:46:17.315 field in general is moving
NOTE Confidence: 0.97644633
00:46:17.535 --> 00:46:19.454 to, like, resolution agnostic, whole
NOTE Confidence: 0.97644633

00:46:19.454 --> 00:46:21.395 slide level, fully context aware
NOTE Confidence: 0.97644633

00:46:21.614 --> 00:46:22.835 aware models where,
NOTE Confidence: 0.9356951

00:46:23.694 --> 00:46:25.430 we would have singular feature
NOTE Confidence: 0.9356951

00:46:25.430 --> 00:46:27.529 representations corresponding the whole slide
NOTE Confidence: 0.9356951

00:46:27.670 --> 00:46:28.789 still capable of doing whole
NOTE Confidence: 0.9356951

00:46:28.789 --> 00:46:30.630 slide level whole slide level
NOTE Confidence: 0.9356951

00:46:30.630 --> 00:46:31.130 tasks.
NOTE Confidence: 0.9447295

00:46:31.509 --> 00:46:32.549 Now the I think your
NOTE Confidence: 0.9447295

00:46:32.549 --> 00:46:34.469 question particularly refers to if
NOTE Confidence: 0.9447295

00:46:34.469 --> 00:46:35.509 you have smaller patches, are
NOTE Confidence: 0.9447295

00:46:35.509 --> 00:46:36.469 you able to do something
NOTE Confidence: 0.9447295

00:46:36.469 --> 00:46:37.635 very fine fine grained? So
NOTE Confidence: 0.9447295

00:46:37.635 --> 00:46:38.675 that's true. So if you
NOTE Confidence: 0.9447295

00:46:38.675 --> 00:46:39.655 have smaller patches,
NOTE Confidence: 0.9642636

00:46:40.195 --> 00:46:41.155 you're able to, you know,
NOTE Confidence: 0.9642636

00:46:41.155 --> 00:46:42.835 separate out tells and and

NOTE Confidence: 0.9642636
00:46:42.835 --> 00:46:44.114 do do specific things. So
NOTE Confidence: 0.9642636
00:46:44.114 --> 00:46:45.315 so what would happen in
NOTE Confidence: 0.9642636
00:46:45.315 --> 00:46:46.275 the long run? I think
NOTE Confidence: 0.9642636
00:46:46.275 --> 00:46:47.635 it would be a combination
NOTE Confidence: 0.9642636
00:46:47.635 --> 00:46:48.594 of the two. So we'll
NOTE Confidence: 0.9642636
00:46:48.594 --> 00:46:49.315 we'll have, like, a whole
NOTE Confidence: 0.9642636
00:46:49.315 --> 00:46:50.810 slide level feature vector that
NOTE Confidence: 0.9642636
00:46:50.810 --> 00:46:51.530 can be used for other
NOTE Confidence: 0.9642636
00:46:51.530 --> 00:46:53.530 downstream tasks, and you have,
NOTE Confidence: 0.9642636
00:46:53.530 --> 00:46:55.370 like, smaller patches that can
NOTE Confidence: 0.9642636
00:46:55.370 --> 00:46:56.570 be used for other kind
NOTE Confidence: 0.9642636
00:46:56.570 --> 00:46:58.270 of more localized region specific
NOTE Confidence: 0.86922824
00:46:58.730 --> 00:47:00.489 region specific tasks. What was
NOTE Confidence: 0.86922824
00:47:00.489 --> 00:47:02.410 your Special state. Uh-huh. Special
NOTE Confidence: 0.86922824
00:47:02.410 --> 00:47:03.530 state. Yes. So the,
NOTE Confidence: 0.9739618

00:47:05.075 --> 00:47:06.035 yeah, majority of the work
NOTE Confidence: 0.9739618

00:47:06.035 --> 00:47:06.935 is based on
NOTE Confidence: 0.9343008

00:47:07.555 --> 00:47:08.755 based on HNE. I completely
NOTE Confidence: 0.9343008

00:47:08.755 --> 00:47:10.195 agree that, you know, as
NOTE Confidence: 0.9343008

00:47:10.195 --> 00:47:11.475 as you have more of
NOTE Confidence: 0.9343008

00:47:11.475 --> 00:47:11.975 this
NOTE Confidence: 0.9302768

00:47:12.435 --> 00:47:14.135 complimentary data, whether it's through
NOTE Confidence: 0.9302768

00:47:14.355 --> 00:47:16.035 special stains or administered chemistry,
NOTE Confidence: 0.9302768

00:47:16.035 --> 00:47:17.415 you can extract more information
NOTE Confidence: 0.9302768

00:47:17.475 --> 00:47:18.760 from these from these images.
NOTE Confidence: 0.9302768

00:47:18.760 --> 00:47:19.840 From the HNE already, we
NOTE Confidence: 0.9302768

00:47:19.840 --> 00:47:20.760 can we can begin to
NOTE Confidence: 0.9302768

00:47:20.760 --> 00:47:22.620 extract more information than than
NOTE Confidence: 0.9302768

00:47:22.680 --> 00:47:23.720 than what we can see.
NOTE Confidence: 0.9302768

00:47:23.720 --> 00:47:25.080 So with by having special
NOTE Confidence: 0.9302768

00:47:25.080 --> 00:47:26.620 stains, you can extract even

NOTE Confidence: 0.9302768
00:47:26.680 --> 00:47:27.500 even more.
NOTE Confidence: 0.99270993
00:47:28.840 --> 00:47:30.360 So in in in the
NOTE Confidence: 0.99270993
00:47:30.360 --> 00:47:31.800 short term, our goal is
NOTE Confidence: 0.99270993
00:47:31.800 --> 00:47:32.460 to see
NOTE Confidence: 0.89152265
00:47:33.915 --> 00:47:35.594 how you're using PathChat. How
NOTE Confidence: 0.89152265
00:47:35.594 --> 00:47:36.415 can we go,
NOTE Confidence: 0.92839944
00:47:37.594 --> 00:47:38.635 from h and e to
NOTE Confidence: 0.92839944
00:47:38.635 --> 00:47:40.315 predicting what special stains or
NOTE Confidence: 0.92839944
00:47:40.315 --> 00:47:41.515 or aminos to chemistry or
NOTE Confidence: 0.92839944
00:47:41.515 --> 00:47:42.715 other ancillary test need need
NOTE Confidence: 0.92839944
00:47:42.715 --> 00:47:44.235 to be ordered, ingest them
NOTE Confidence: 0.92839944
00:47:44.235 --> 00:47:45.455 into the same context,
NOTE Confidence: 0.98468715
00:47:45.835 --> 00:47:46.715 and see if the model
NOTE Confidence: 0.98468715
00:47:46.715 --> 00:47:47.594 can get to get to
NOTE Confidence: 0.98468715
00:47:47.594 --> 00:47:48.815 a pathology report.
NOTE Confidence: 0.9658531

00:47:51.000 --> 00:47:52.200 It's it's a little bit
NOTE Confidence: 0.9658531

00:47:52.200 --> 00:47:53.320 difficult to go to, like,
NOTE Confidence: 0.9658531

00:47:53.320 --> 00:47:54.840 oral outcome using some of
NOTE Confidence: 0.9658531

00:47:54.840 --> 00:47:56.280 these special stains if there's
NOTE Confidence: 0.9658531

00:47:56.280 --> 00:47:57.560 not large enough cohorts available.
NOTE Confidence: 0.9658531

00:47:57.560 --> 00:47:58.680 But if they're available, that's
NOTE Confidence: 0.9658531

00:47:58.680 --> 00:47:59.580 very much possible.
NOTE Confidence: 0.6847708

00:48:00.040 --> 00:48:01.239 Just to give an example
NOTE Confidence: 0.6847708

00:48:01.239 --> 00:48:02.040 at the bottom of the
NOTE Confidence: 0.6847708

00:48:02.040 --> 00:48:03.020 special stain.
NOTE Confidence: 0.74119866

00:48:03.474 --> 00:48:04.675 We use Brightcove for five
NOTE Confidence: 0.74119866

00:48:04.675 --> 00:48:06.275 percent study. Just generally, for
NOTE Confidence: 0.74119866

00:48:06.275 --> 00:48:07.315 us, it's five percent for
NOTE Confidence: 0.74119866

00:48:07.315 --> 00:48:08.594 five percent and light and
NOTE Confidence: 0.74119866

00:48:08.594 --> 00:48:09.815 dark end of the aesthetic.
NOTE Confidence: 0.9748

00:48:10.194 --> 00:48:12.535 But using, this algorithm,

NOTE Confidence: 0.9387533
00:48:13.635 --> 00:48:15.734 one of the AI models
NOTE Confidence: 0.33313757
00:48:16.194 --> 00:48:17.015 was working
NOTE Confidence: 0.81688935
00:48:33.344 --> 00:48:33.984 Mhmm. So that's what I'm
NOTE Confidence: 0.81688935
00:48:33.984 --> 00:48:35.185 talking about. We really don't
NOTE Confidence: 0.81688935
00:48:35.185 --> 00:48:36.065 use it that way, but
NOTE Confidence: 0.81688935
00:48:36.145 --> 00:48:37.505 Yeah. Maybe the AI can't
NOTE Confidence: 0.81688935
00:48:37.505 --> 00:48:38.705 figure out those set up
NOTE Confidence: 0.81688935
00:48:38.705 --> 00:48:39.825 on your side of us.
NOTE Confidence: 0.81688935
00:48:39.825 --> 00:48:41.185 Yeah. They they they possibly
NOTE Confidence: 0.81688935
00:48:41.185 --> 00:48:42.705 used handcrafted features, like, four
NOTE Confidence: 0.81688935
00:48:42.705 --> 00:48:44.224 hundred different handcrafted features and
NOTE Confidence: 0.81688935
00:48:44.224 --> 00:48:45.585 then correlate them with the
NOTE Confidence: 0.81688935
00:48:45.744 --> 00:48:47.105 with outcome. You could potentially
NOTE Confidence: 0.81688935
00:48:47.105 --> 00:48:48.600 also do that directly, right,
NOTE Confidence: 0.81688935
00:48:48.680 --> 00:48:49.180 using,
NOTE Confidence: 0.9759318

00:48:49.560 --> 00:48:50.620 like, deep features.
NOTE Confidence: 0.95314705
00:48:51.560 --> 00:48:52.060 Yeah.
NOTE Confidence: 0.47777256
00:48:56.360 --> 00:48:57.260 Question over here.
NOTE Confidence: 0.98958015
00:48:57.640 --> 00:48:58.140 Yes.
NOTE Confidence: 0.7923998
00:49:01.364 --> 00:49:02.025 Go ahead.
NOTE Confidence: 0.95844257
00:49:03.125 --> 00:49:04.825 So pathologists have, like,
NOTE Confidence: 0.8465375
00:49:05.205 --> 00:49:06.725 basic textbooks that say this
NOTE Confidence: 0.8465375
00:49:06.725 --> 00:49:07.525 is this disease and this
NOTE Confidence: 0.8465375
00:49:07.525 --> 00:49:08.565 is not this disease. Have
NOTE Confidence: 0.8465375
00:49:08.565 --> 00:49:09.765 you gone back to the
NOTE Confidence: 0.8465375
00:49:09.765 --> 00:49:11.605 features that your networks are
NOTE Confidence: 0.8465375
00:49:11.605 --> 00:49:13.125 picking out and say, we're
NOTE Confidence: 0.8465375
00:49:13.125 --> 00:49:14.325 picking out the same features
NOTE Confidence: 0.8465375
00:49:14.325 --> 00:49:15.420 that sort of biologists are
NOTE Confidence: 0.8465375
00:49:15.420 --> 00:49:17.359 using for characterizing these diseases.
NOTE Confidence: 0.7247029
00:49:18.380 --> 00:49:19.180 You know, you showed some

NOTE Confidence: 0.7247029
00:49:19.180 --> 00:49:20.460 key maps that kinda show
NOTE Confidence: 0.7247029
00:49:20.460 --> 00:49:22.480 broad regions. But Yeah. Certainly,
NOTE Confidence: 0.7247029
00:49:22.619 --> 00:49:23.739 some disease are defined by
NOTE Confidence: 0.68994576
00:49:25.020 --> 00:49:26.880 that mnemonic cell types versus
NOTE Confidence: 0.68994576
00:49:26.940 --> 00:49:27.440 some.
NOTE Confidence: 0.96825933
00:49:29.055 --> 00:49:30.094 Yeah. Yeah. So a lot
NOTE Confidence: 0.96825933
00:49:30.094 --> 00:49:31.295 of the disease specific work
NOTE Confidence: 0.96825933
00:49:31.295 --> 00:49:32.335 that we have done, we
NOTE Confidence: 0.96825933
00:49:32.335 --> 00:49:33.535 we did do that. We,
NOTE Confidence: 0.9300266
00:49:34.335 --> 00:49:35.695 asked pathologists to look at
NOTE Confidence: 0.9300266
00:49:35.695 --> 00:49:37.395 the high tension regions,
NOTE Confidence: 0.916184
00:49:39.055 --> 00:49:41.055 and just narrate what they
NOTE Confidence: 0.916184
00:49:41.055 --> 00:49:41.714 were seeing.
NOTE Confidence: 0.92842144
00:49:42.590 --> 00:49:43.390 And we have some of
NOTE Confidence: 0.92842144
00:49:43.390 --> 00:49:44.830 that analysis. In in in
NOTE Confidence: 0.92842144

00:49:44.830 --> 00:49:45.790 a few studies, we also
NOTE Confidence: 0.92842144

00:49:45.790 --> 00:49:47.430 did quantitative analysis on top
NOTE Confidence: 0.92842144

00:49:47.430 --> 00:49:48.910 of it. So high attention
NOTE Confidence: 0.92842144

00:49:48.910 --> 00:49:50.989 regions, and then you basically
NOTE Confidence: 0.92842144

00:49:50.989 --> 00:49:52.530 quantify all the
NOTE Confidence: 0.9753863

00:49:52.910 --> 00:49:54.690 all the cells, classify them,
NOTE Confidence: 0.9736047

00:49:55.075 --> 00:49:57.075 extract handcrafted features, and correlate
NOTE Confidence: 0.9736047

00:49:57.075 --> 00:49:58.995 those handcrafted features with the
NOTE Confidence: 0.9736047

00:49:58.995 --> 00:50:00.435 deep the the the the
NOTE Confidence: 0.9736047

00:50:00.435 --> 00:50:01.635 regions that the deep model
NOTE Confidence: 0.9736047

00:50:01.635 --> 00:50:02.614 was essentially using
NOTE Confidence: 0.9526822

00:50:02.915 --> 00:50:04.035 to get a more quantitative
NOTE Confidence: 0.9526822

00:50:04.035 --> 00:50:05.315 assessment of that. Because this
NOTE Confidence: 0.9526822

00:50:05.315 --> 00:50:06.515 was done for, like, cancer
NOTE Confidence: 0.9526822

00:50:06.515 --> 00:50:08.289 of unknown primary thing or
NOTE Confidence: 0.9526822

00:50:08.289 --> 00:50:09.410 also for the cardiac,

NOTE Confidence: 0.96643066
00:50:09.809 --> 00:50:11.430 allograft biopsy study.
NOTE Confidence: 0.99683654
00:50:12.529 --> 00:50:13.170 It was done for a
NOTE Confidence: 0.99683654
00:50:13.170 --> 00:50:14.309 number of studies. Yeah.
NOTE Confidence: 0.8302023
00:50:14.690 --> 00:50:15.809 Scale of those regions, is
NOTE Confidence: 0.8302023
00:50:15.809 --> 00:50:16.210 it,
NOTE Confidence: 0.9093778
00:50:17.969 --> 00:50:19.489 the cell level scale? Was
NOTE Confidence: 0.9093778
00:50:19.489 --> 00:50:21.494 it sort of a larger,
NOTE Confidence: 0.9093778
00:50:21.555 --> 00:50:22.994 like, tens of cells, hundreds
NOTE Confidence: 0.9093778
00:50:22.994 --> 00:50:24.515 of cells? Yeah. It depends
NOTE Confidence: 0.9093778
00:50:24.515 --> 00:50:25.895 on the study. So,
NOTE Confidence: 0.9126366
00:50:26.755 --> 00:50:27.875 for example, for the cancer
NOTE Confidence: 0.9126366
00:50:27.875 --> 00:50:29.075 found on primary study, we
NOTE Confidence: 0.9126366
00:50:29.075 --> 00:50:30.355 were able to do that
NOTE Confidence: 0.9126366
00:50:30.355 --> 00:50:31.315 at a cell level and
NOTE Confidence: 0.9126366
00:50:31.315 --> 00:50:32.594 then get a quantitative assessment
NOTE Confidence: 0.9126366

00:50:32.594 --> 00:50:33.715 that the model is predominantly

NOTE Confidence: 0.9126366

00:50:33.715 --> 00:50:35.109 looking at at tumor regions

NOTE Confidence: 0.9126366

00:50:35.109 --> 00:50:36.390 and then what what handcrafted

NOTE Confidence: 0.9126366

00:50:36.390 --> 00:50:38.090 features were being used. Yeah.

NOTE Confidence: 0.7834573

00:50:41.270 --> 00:50:41.770 Yeah.

NOTE Confidence: 0.71318436

00:50:47.065 --> 00:50:47.565 Yes.

NOTE Confidence: 0.7476492

00:50:51.545 --> 00:50:52.045 Yes.

NOTE Confidence: 0.5444502

00:50:56.105 --> 00:50:57.545 Trading and our councils that

NOTE Confidence: 0.5444502

00:50:57.545 --> 00:50:58.045 are

NOTE Confidence: 0.8265823

00:50:58.665 --> 00:51:01.145 very rare. Yes. Yes. Every

NOTE Confidence: 0.8265823

00:51:01.145 --> 00:51:03.130 video. Right? Right. So what

NOTE Confidence: 0.8265823

00:51:03.130 --> 00:51:04.090 what is the future of

NOTE Confidence: 0.8265823

00:51:04.090 --> 00:51:05.450 those concepts in terms of

NOTE Confidence: 0.99399495

00:51:06.650 --> 00:51:07.870 Yeah. Yeah. So

NOTE Confidence: 0.92447805

00:51:08.570 --> 00:51:10.090 we try so so so

NOTE Confidence: 0.92447805

00:51:10.090 --> 00:51:10.590 the,

NOTE Confidence: 0.9223956
00:51:12.570 --> 00:51:13.170 there there are a number
NOTE Confidence: 0.9223956
00:51:13.170 --> 00:51:14.489 of different datasets used in
NOTE Confidence: 0.9223956
00:51:14.489 --> 00:51:15.515 in in this process. Right?
NOTE Confidence: 0.9223956
00:51:15.515 --> 00:51:16.315 So the for the for
NOTE Confidence: 0.9223956
00:51:16.315 --> 00:51:17.675 the pre training data, we're
NOTE Confidence: 0.9223956
00:51:17.675 --> 00:51:18.715 just using images. There's a
NOTE Confidence: 0.9223956
00:51:18.715 --> 00:51:19.614 huge disparity.
NOTE Confidence: 0.9719125
00:51:20.955 --> 00:51:22.155 But for the instruction dataset,
NOTE Confidence: 0.9719125
00:51:22.155 --> 00:51:23.195 there isn't a huge disparity
NOTE Confidence: 0.9719125
00:51:23.195 --> 00:51:24.475 because we try to maximize
NOTE Confidence: 0.9719125
00:51:24.475 --> 00:51:25.594 for diversity, and it's very
NOTE Confidence: 0.9719125
00:51:25.594 --> 00:51:27.114 difficult to collect collect that
NOTE Confidence: 0.9719125
00:51:27.114 --> 00:51:27.614 data.
NOTE Confidence: 0.5751502
00:51:28.820 --> 00:51:29.320 But
NOTE Confidence: 0.97118187
00:51:29.860 --> 00:51:31.140 the the performance is obviously
NOTE Confidence: 0.97118187

00:51:31.140 --> 00:51:32.920 much better on common,
NOTE Confidence: 0.96019983

00:51:33.700 --> 00:51:35.060 common disease entities and is
NOTE Confidence: 0.96019983

00:51:35.060 --> 00:51:35.860 not so much,
NOTE Confidence: 0.9931581

00:51:36.580 --> 00:51:37.540 not not so much so
NOTE Confidence: 0.9931581

00:51:37.540 --> 00:51:38.760 on the rarer entities.
NOTE Confidence: 0.97593945

00:51:40.580 --> 00:51:40.980 It's
NOTE Confidence: 0.92242765

00:51:41.380 --> 00:51:43.265 it it it's possible
NOTE Confidence: 0.9124344

00:51:43.805 --> 00:51:45.005 to use few shot learning
NOTE Confidence: 0.9124344

00:51:45.005 --> 00:51:45.885 and everything that we're doing
NOTE Confidence: 0.9124344

00:51:45.885 --> 00:51:46.684 in terms of which which
NOTE Confidence: 0.9124344

00:51:46.684 --> 00:51:48.605 feature representation to improve performance
NOTE Confidence: 0.9124344

00:51:48.605 --> 00:51:50.145 of these on these entities.
NOTE Confidence: 0.99814767

00:51:50.765 --> 00:51:52.285 The issue is that it's
NOTE Confidence: 0.99814767

00:51:52.285 --> 00:51:53.905 very difficult to validate it.
NOTE Confidence: 0.98782206

00:51:54.980 --> 00:51:56.260 Would you trust an algorithm
NOTE Confidence: 0.98782206

00:51:56.260 --> 00:51:57.640 that was trained on two,

NOTE Confidence: 0.9582329
00:51:58.099 --> 00:51:59.779 two images and was evaluated
NOTE Confidence: 0.9582329
00:51:59.779 --> 00:52:01.539 on five? Right. So so
NOTE Confidence: 0.9582329
00:52:01.539 --> 00:52:02.579 so the issue is not
NOTE Confidence: 0.9582329
00:52:02.579 --> 00:52:03.619 around. So so I trust
NOTE Confidence: 0.9582329
00:52:03.619 --> 00:52:04.739 few shot learning because I
NOTE Confidence: 0.9582329
00:52:04.739 --> 00:52:05.539 can see it works so
NOTE Confidence: 0.9582329
00:52:05.539 --> 00:52:07.140 well, and it's in line
NOTE Confidence: 0.9582329
00:52:07.140 --> 00:52:08.020 with what the rest of
NOTE Confidence: 0.9582329
00:52:08.020 --> 00:52:09.575 the machine community is thinking
NOTE Confidence: 0.9582329
00:52:09.575 --> 00:52:10.714 and every other field.
NOTE Confidence: 0.98585624
00:52:11.575 --> 00:52:12.375 So it should work here
NOTE Confidence: 0.98585624
00:52:12.375 --> 00:52:14.855 too. But the reason I
NOTE Confidence: 0.98585624
00:52:14.855 --> 00:52:15.895 would not trust it is
NOTE Confidence: 0.98585624
00:52:15.895 --> 00:52:17.255 because there isn't enough data
NOTE Confidence: 0.98585624
00:52:17.255 --> 00:52:18.934 to validate it. So if
NOTE Confidence: 0.98585624

00:52:18.934 --> 00:52:20.055 if a rare entity has
NOTE Confidence: 0.98585624

00:52:20.055 --> 00:52:21.594 ten cases, twelve cases,
NOTE Confidence: 0.8904331

00:52:22.089 --> 00:52:22.890 and and the model is
NOTE Confidence: 0.8904331

00:52:22.890 --> 00:52:23.930 fine on all all twelve
NOTE Confidence: 0.8904331

00:52:23.930 --> 00:52:24.969 of them, should we now
NOTE Confidence: 0.8904331

00:52:24.969 --> 00:52:25.869 trust it or
NOTE Confidence: 0.9427823

00:52:26.250 --> 00:52:27.210 or should we still get
NOTE Confidence: 0.9427823

00:52:27.210 --> 00:52:28.270 more data? So,
NOTE Confidence: 0.9974048

00:52:29.130 --> 00:52:30.670 the issue is around validation
NOTE Confidence: 0.9974048

00:52:30.730 --> 00:52:31.630 for rare diseases.
NOTE Confidence: 0.75189775

00:52:32.650 --> 00:52:34.010 One more question. You showed
NOTE Confidence: 0.75189775

00:52:34.010 --> 00:52:34.449 us in this last part
NOTE Confidence: 0.75189775

00:52:34.449 --> 00:52:35.319 of the the talk is
NOTE Confidence: 0.7378987

00:52:51.000 --> 00:52:53.100 Yes. Of the cells. Yeah.
NOTE Confidence: 0.996685

00:52:53.719 --> 00:52:55.020 Yeah. So we we
NOTE Confidence: 0.9757311

00:52:55.320 --> 00:52:56.120 have tried a number of

NOTE Confidence: 0.9757311

00:52:56.120 --> 00:52:58.600 different approaches, including cyclic approaches.

NOTE Confidence: 0.9757311

00:52:58.600 --> 00:52:59.719 So cyclic approach is very

NOTE Confidence: 0.9757311

00:52:59.719 --> 00:53:00.840 common in computer science where

NOTE Confidence: 0.9757311

00:53:00.840 --> 00:53:02.219 you go from one,

NOTE Confidence: 0.96978176

00:53:02.920 --> 00:53:03.960 modality to the other and

NOTE Confidence: 0.96978176

00:53:03.960 --> 00:53:05.100 then from the other modality

NOTE Confidence: 0.96978176

00:53:05.320 --> 00:53:06.675 back, right, and hope that

NOTE Confidence: 0.96978176

00:53:06.675 --> 00:53:08.275 this cyclic approach would would

NOTE Confidence: 0.96978176

00:53:08.275 --> 00:53:09.155 would approve it. So it

NOTE Confidence: 0.96978176

00:53:09.155 --> 00:53:10.515 is possible to make that

NOTE Confidence: 0.96978176

00:53:10.515 --> 00:53:11.875 prediction. In particular, if you

NOTE Confidence: 0.96978176

00:53:11.875 --> 00:53:12.375 use

NOTE Confidence: 0.9908631

00:53:12.675 --> 00:53:14.195 a self supervised model that

NOTE Confidence: 0.9908631

00:53:14.195 --> 00:53:14.695 that

NOTE Confidence: 0.96427953

00:53:15.075 --> 00:53:16.835 is used to predicting pathology

NOTE Confidence: 0.96427953

00:53:16.835 --> 00:53:18.915 images, like, just just just
NOTE Confidence: 0.96427953

00:53:18.915 --> 00:53:20.640 predicting missing patches or something.
NOTE Confidence: 0.9921228

00:53:21.200 --> 00:53:22.239 So it's possible to go
NOTE Confidence: 0.9921228

00:53:22.239 --> 00:53:23.299 the other way as well.
NOTE Confidence: 0.97843015

00:53:24.480 --> 00:53:24.980 Yeah.
NOTE Confidence: 0.8109032

00:53:25.599 --> 00:53:26.099 Yes.
NOTE Confidence: 0.39963758

00:53:26.960 --> 00:53:27.619 Two part.
NOTE Confidence: 0.5877898

00:53:28.079 --> 00:53:29.140 Looking over, like,
NOTE Confidence: 0.8373222

00:53:30.400 --> 00:53:31.140 one thing.
NOTE Confidence: 0.94843185

00:53:32.160 --> 00:53:33.619 These, you know, images,
NOTE Confidence: 0.7339425

00:53:34.079 --> 00:53:35.875 these large image databases are
NOTE Confidence: 0.7339425

00:53:35.875 --> 00:53:36.375 based
NOTE Confidence: 0.75690174

00:53:36.835 --> 00:53:38.675 on years off decades worth
NOTE Confidence: 0.75690174

00:53:38.675 --> 00:53:39.175 of
NOTE Confidence: 0.7955534

00:53:39.555 --> 00:53:40.695 of of data with associated
NOTE Confidence: 0.7955534

00:53:40.915 --> 00:53:41.415 diagnoses.

NOTE Confidence: 0.72448283
00:53:42.035 --> 00:53:43.155 But anyone who's got a
NOTE Confidence: 0.72448283
00:53:43.155 --> 00:53:44.275 valve for more than five
NOTE Confidence: 0.72448283
00:53:44.275 --> 00:53:45.735 years realizes that
NOTE Confidence: 0.93358177
00:53:46.035 --> 00:53:47.095 that our diagnostics,
NOTE Confidence: 0.7328665
00:53:48.160 --> 00:53:49.520 you know, change over time.
NOTE Confidence: 0.7328665
00:53:49.520 --> 00:53:50.719 Right? And so so what
NOTE Confidence: 0.7328665
00:53:50.800 --> 00:53:51.920 you know, in years of
NOTE Confidence: 0.7328665
00:53:51.920 --> 00:53:53.360 quality, we've been called bleeding
NOTE Confidence: 0.7328665
00:53:53.360 --> 00:53:54.580 by recipients. It doesn't
NOTE Confidence: 0.9060521
00:53:55.440 --> 00:53:57.200 exist in. Yeah. So so
NOTE Confidence: 0.9060521
00:53:57.200 --> 00:53:57.860 how do
NOTE Confidence: 0.95659405
00:53:58.320 --> 00:53:59.540 you how do you expunge
NOTE Confidence: 0.95659405
00:53:59.760 --> 00:54:01.219 those sort of Yeah.
NOTE Confidence: 0.5417275
00:54:01.600 --> 00:54:02.980 Things from from the
NOTE Confidence: 0.69917226
00:54:03.305 --> 00:54:05.645 train models Right. Is. Yeah.
NOTE Confidence: 0.8698166

00:54:06.905 --> 00:54:08.265 So I'll I'll give that
NOTE Confidence: 0.8698166

00:54:08.265 --> 00:54:09.465 person a second. Right. That
NOTE Confidence: 0.8698166

00:54:09.465 --> 00:54:10.425 that that's a really good
NOTE Confidence: 0.8698166

00:54:10.425 --> 00:54:11.625 question. And we had a
NOTE Confidence: 0.8698166

00:54:11.625 --> 00:54:12.665 huge problem with this when
NOTE Confidence: 0.8698166

00:54:12.665 --> 00:54:13.705 we were con constructing the
NOTE Confidence: 0.8698166

00:54:13.705 --> 00:54:14.685 instruction dataset.
NOTE Confidence: 0.9109695

00:54:17.130 --> 00:54:17.869 The the
NOTE Confidence: 0.91927326

00:54:18.730 --> 00:54:19.690 the the solution we came
NOTE Confidence: 0.91927326

00:54:19.690 --> 00:54:20.589 up with was, obviously,
NOTE Confidence: 0.8247763

00:54:21.210 --> 00:54:22.489 to just manually evaluate each
NOTE Confidence: 0.8247763

00:54:22.489 --> 00:54:24.510 and every, data training point.
NOTE Confidence: 0.99079067

00:54:25.049 --> 00:54:26.029 But since then,
NOTE Confidence: 0.9861918

00:54:26.969 --> 00:54:28.089 we have come up with,
NOTE Confidence: 0.9861918

00:54:28.089 --> 00:54:28.589 like,
NOTE Confidence: 0.98093545

00:54:29.484 --> 00:54:31.005 equivalency map for what what

NOTE Confidence: 0.98093545
00:54:31.005 --> 00:54:32.045 a certain entity used to
NOTE Confidence: 0.98093545
00:54:32.045 --> 00:54:32.844 be called and how it
NOTE Confidence: 0.98093545
00:54:32.844 --> 00:54:33.344 merged.
NOTE Confidence: 0.99309033
00:54:33.964 --> 00:54:35.165 That has helped us clean
NOTE Confidence: 0.99309033
00:54:35.165 --> 00:54:36.464 up a lot of the,
NOTE Confidence: 0.93265057
00:54:37.244 --> 00:54:38.925 old data. And we're also
NOTE Confidence: 0.93265057
00:54:38.925 --> 00:54:40.045 actively thinking about what would
NOTE Confidence: 0.93265057
00:54:40.045 --> 00:54:41.325 happen in the future because,
NOTE Confidence: 0.93265057
00:54:41.325 --> 00:54:41.964 you know, you have a
NOTE Confidence: 0.93265057
00:54:41.964 --> 00:54:43.950 multimodal large language model, which
NOTE Confidence: 0.93265057
00:54:43.950 --> 00:54:45.030 will be when we spent
NOTE Confidence: 0.93265057
00:54:45.030 --> 00:54:46.730 a significant amount of computational
NOTE Confidence: 0.9710913
00:54:47.030 --> 00:54:48.469 resource to train it, we
NOTE Confidence: 0.9710913
00:54:48.469 --> 00:54:49.670 don't wanna train it when
NOTE Confidence: 0.9710913
00:54:49.670 --> 00:54:50.710 there's a new blue book.
NOTE Confidence: 0.9710913

00:54:50.710 --> 00:54:52.250 Right? So so,
NOTE Confidence: 0.9803887

00:54:52.790 --> 00:54:54.150 we're looking into, like, retrieval
NOTE Confidence: 0.9803887

00:54:54.150 --> 00:54:55.910 augmented generation, other techniques that
NOTE Confidence: 0.9803887

00:54:55.910 --> 00:54:56.730 can be used,
NOTE Confidence: 0.945373

00:54:57.910 --> 00:55:00.045 and leverage to to update
NOTE Confidence: 0.945373

00:55:00.045 --> 00:55:01.105 some of those diagnoses.
NOTE Confidence: 0.83720845

00:55:01.405 --> 00:55:02.765 Okay. And then the other
NOTE Confidence: 0.83720845

00:55:02.765 --> 00:55:03.885 one about you talked briefly
NOTE Confidence: 0.83720845

00:55:03.885 --> 00:55:04.785 about the biases
NOTE Confidence: 0.7852704

00:55:05.165 --> 00:55:05.905 in the model,
NOTE Confidence: 0.85549563

00:55:06.364 --> 00:55:07.885 up against certain, you know,
NOTE Confidence: 0.85549563

00:55:07.885 --> 00:55:08.785 patient populations,
NOTE Confidence: 0.90041333

00:55:09.325 --> 00:55:10.525 you know, ethnic gender, those
NOTE Confidence: 0.90041333

00:55:10.525 --> 00:55:12.160 sort of things. And and
NOTE Confidence: 0.90041333

00:55:12.160 --> 00:55:13.040 I know that there's a
NOTE Confidence: 0.90041333

00:55:13.040 --> 00:55:14.239 lot of of work that

NOTE Confidence: 0.90041333
00:55:14.239 --> 00:55:15.600 I'm trying to, like okay.
NOTE Confidence: 0.90041333
00:55:15.600 --> 00:55:16.960 You recognize that there's a
NOTE Confidence: 0.90041333
00:55:16.960 --> 00:55:18.320 a bias in the model,
NOTE Confidence: 0.90041333
00:55:18.320 --> 00:55:19.360 and so you try to
NOTE Confidence: 0.90041333
00:55:19.360 --> 00:55:19.860 tweak
NOTE Confidence: 0.9976622
00:55:20.239 --> 00:55:21.860 the model a little bit
NOTE Confidence: 0.9976622
00:55:22.080 --> 00:55:23.360 to remove some of that
NOTE Confidence: 0.9976622
00:55:23.360 --> 00:55:23.860 bias.
NOTE Confidence: 0.9211429
00:55:25.085 --> 00:55:26.204 And what I've always been
NOTE Confidence: 0.9211429
00:55:26.204 --> 00:55:27.484 surprised about is why do
NOTE Confidence: 0.9211429
00:55:27.484 --> 00:55:29.565 people simply not pursue a
NOTE Confidence: 0.9211429
00:55:29.565 --> 00:55:30.765 different course of action to
NOTE Confidence: 0.9211429
00:55:30.765 --> 00:55:32.045 say that, you know, for
NOTE Confidence: 0.9211429
00:55:32.045 --> 00:55:33.164 example, we need a different
NOTE Confidence: 0.9211429
00:55:33.164 --> 00:55:34.444 model for men than we
NOTE Confidence: 0.9211429

00:55:34.444 --> 00:55:35.184 do for women,
NOTE Confidence: 0.9747786

00:55:35.645 --> 00:55:36.684 rather than trying to come
NOTE Confidence: 0.9747786

00:55:36.684 --> 00:55:37.964 up with one model that
NOTE Confidence: 0.9747786

00:55:37.964 --> 00:55:39.025 works for both.
NOTE Confidence: 0.8776072

00:55:40.080 --> 00:55:41.220 And I wonder what the
NOTE Confidence: 0.8776072

00:55:41.440 --> 00:55:43.440 the base is. Right. That's
NOTE Confidence: 0.8776072

00:55:43.440 --> 00:55:44.660 a great question. I
NOTE Confidence: 0.95171267

00:55:45.360 --> 00:55:46.239 I I I think a
NOTE Confidence: 0.95171267

00:55:46.239 --> 00:55:47.120 lot of it has to
NOTE Confidence: 0.95171267

00:55:47.120 --> 00:55:47.940 do with data.
NOTE Confidence: 0.99381113

00:55:49.040 --> 00:55:50.420 How much data is available,
NOTE Confidence: 0.8760202

00:55:51.120 --> 00:55:51.860 for this?
NOTE Confidence: 0.9573753

00:55:53.285 --> 00:55:54.965 Would a model if there's
NOTE Confidence: 0.9573753

00:55:54.965 --> 00:55:56.885 disparity between men and women
NOTE Confidence: 0.9573753

00:55:56.885 --> 00:55:57.705 for a model,
NOTE Confidence: 0.8792491

00:55:58.885 --> 00:56:00.425 will training separate models

NOTE Confidence: 0.98192775

00:56:00.885 --> 00:56:02.345 help? It could be.

NOTE Confidence: 0.95417225

00:56:03.605 --> 00:56:04.105 But

NOTE Confidence: 0.9846539

00:56:04.550 --> 00:56:06.010 if there are no known

NOTE Confidence: 0.9846539

00:56:06.230 --> 00:56:07.369 morphologic differences,

NOTE Confidence: 0.9730071

00:56:09.109 --> 00:56:10.230 there are there's a high

NOTE Confidence: 0.9730071

00:56:10.230 --> 00:56:11.270 chance that there could be

NOTE Confidence: 0.9730071

00:56:11.270 --> 00:56:12.230 some other reason. And let

NOTE Confidence: 0.9730071

00:56:12.230 --> 00:56:13.290 me give you an example.

NOTE Confidence: 0.9965954

00:56:14.469 --> 00:56:15.130 So we

NOTE Confidence: 0.9600919

00:56:15.750 --> 00:56:17.285 looked at, you know, a

NOTE Confidence: 0.9600919

00:56:17.285 --> 00:56:18.165 a lot of these models,

NOTE Confidence: 0.9600919

00:56:18.165 --> 00:56:19.765 very fundamental tasks. Can you

NOTE Confidence: 0.9600919

00:56:19.765 --> 00:56:21.625 subtype, you know, breast carcinoma?

NOTE Confidence: 0.9600919

00:56:21.685 --> 00:56:22.645 And can you subtype non

NOTE Confidence: 0.9600919

00:56:22.645 --> 00:56:23.864 small cell lung cancer?

NOTE Confidence: 0.9402237

00:56:24.405 --> 00:56:25.685 These are these are tasks
NOTE Confidence: 0.9402237

00:56:25.685 --> 00:56:26.805 that have been essentially solved
NOTE Confidence: 0.9402237

00:56:26.805 --> 00:56:27.765 by machine learning, where you
NOTE Confidence: 0.9402237

00:56:27.765 --> 00:56:28.505 can get
NOTE Confidence: 0.9614808

00:56:29.010 --> 00:56:30.050 models with a zero point
NOTE Confidence: 0.9614808

00:56:30.050 --> 00:56:32.130 nine nine AUC, and they're
NOTE Confidence: 0.9614808

00:56:32.130 --> 00:56:33.030 perfect models.
NOTE Confidence: 0.9124246

00:56:34.370 --> 00:56:36.130 And then we apply them
NOTE Confidence: 0.9124246

00:56:36.130 --> 00:56:36.630 to
NOTE Confidence: 0.9762641

00:56:37.090 --> 00:56:38.930 data from MGH and the
NOTE Confidence: 0.9762641

00:56:38.930 --> 00:56:40.310 and the Brigham and stratified
NOTE Confidence: 0.99187976

00:56:40.930 --> 00:56:42.070 by, you know,
NOTE Confidence: 0.86438614

00:56:42.585 --> 00:56:44.638 protected subgroups. And we you
NOTE Confidence: 0.86438614

00:56:44.638 --> 00:56:45.625 you you you find that
NOTE Confidence: 0.86438614

00:56:45.625 --> 00:56:46.445 there are these
NOTE Confidence: 0.9963276

00:56:46.984 --> 00:56:48.665 large disparities across some of

NOTE Confidence: 0.9963276

00:56:48.665 --> 00:56:49.325 these groups.

NOTE Confidence: 0.98565465

00:56:49.785 --> 00:56:50.265 And,

NOTE Confidence: 0.9601769

00:56:50.585 --> 00:56:51.945 then you start looking deeper.

NOTE Confidence: 0.9601769

00:56:51.945 --> 00:56:53.224 Like, why is this happening?

NOTE Confidence: 0.9601769

00:56:53.224 --> 00:56:53.724 And

NOTE Confidence: 0.96278906

00:56:55.150 --> 00:56:56.110 is is there something else,

NOTE Confidence: 0.96278906

00:56:56.110 --> 00:56:57.630 some other confounding variable that's

NOTE Confidence: 0.96278906

00:56:57.630 --> 00:56:59.230 contributing to this? And you

NOTE Confidence: 0.96278906

00:56:59.230 --> 00:57:01.170 suddenly find out that, well,

NOTE Confidence: 0.9493439

00:57:01.469 --> 00:57:02.830 patients who don't have,

NOTE Confidence: 0.9791982

00:57:03.790 --> 00:57:05.330 insurance tend to get diagnosed

NOTE Confidence: 0.9791982

00:57:05.390 --> 00:57:06.590 late. And there are not

NOTE Confidence: 0.9791982

00:57:06.590 --> 00:57:07.950 enough advanced cases in your

NOTE Confidence: 0.9791982

00:57:07.950 --> 00:57:08.775 training set,

NOTE Confidence: 0.9996222

00:57:09.895 --> 00:57:10.395 because

NOTE Confidence: 0.9502012

00:57:11.255 --> 00:57:12.454 the the training data came
NOTE Confidence: 0.9502012

00:57:12.454 --> 00:57:13.734 from a medical center where
NOTE Confidence: 0.9502012

00:57:13.734 --> 00:57:15.335 patients, you know, was in
NOTE Confidence: 0.9502012

00:57:15.335 --> 00:57:16.694 a region where most patients
NOTE Confidence: 0.9502012

00:57:16.694 --> 00:57:18.635 had insurance and were diagnosed
NOTE Confidence: 0.9502012

00:57:18.694 --> 00:57:20.295 early. And, there are lots
NOTE Confidence: 0.9502012

00:57:20.295 --> 00:57:20.535 of,
NOTE Confidence: 0.9811688

00:57:21.335 --> 00:57:22.375 early cases in that in
NOTE Confidence: 0.9811688

00:57:22.375 --> 00:57:23.494 that cohort, but not enough
NOTE Confidence: 0.9811688

00:57:23.494 --> 00:57:24.450 advanced cases.
NOTE Confidence: 0.89386946

00:57:24.930 --> 00:57:26.070 This is just one example.
NOTE Confidence: 0.89386946

00:57:26.290 --> 00:57:27.650 So often the confounding reason
NOTE Confidence: 0.89386946

00:57:27.650 --> 00:57:28.690 for why these why these
NOTE Confidence: 0.89386946

00:57:28.690 --> 00:57:31.030 disparities exist is just completely,
NOTE Confidence: 0.89386946

00:57:31.089 --> 00:57:31.910 completely different.
NOTE Confidence: 0.8323152

00:57:32.369 --> 00:57:33.170 It might be interesting to

NOTE Confidence: 0.8323152
00:57:33.170 --> 00:57:34.130 do the other way around
NOTE Confidence: 0.8323152
00:57:34.130 --> 00:57:34.869 and see if
NOTE Confidence: 0.84113336
00:57:35.170 --> 00:57:36.530 you could show the the
NOTE Confidence: 0.84113336
00:57:36.530 --> 00:57:37.890 bowel cone cancer and have
NOTE Confidence: 0.84113336
00:57:37.890 --> 00:57:38.930 it predict whether it was
NOTE Confidence: 0.84113336
00:57:38.930 --> 00:57:40.695 from the Right. Example. I
NOTE Confidence: 0.84113336
00:57:40.695 --> 00:57:41.975 didn't use Yeah. To see
NOTE Confidence: 0.84113336
00:57:41.975 --> 00:57:43.655 if there are effects. Right.
NOTE Confidence: 0.84113336
00:57:43.655 --> 00:57:45.095 There there there there are
NOTE Confidence: 0.84113336
00:57:45.095 --> 00:57:46.135 a number of studies that
NOTE Confidence: 0.84113336
00:57:46.135 --> 00:57:47.495 have shown this where they've
NOTE Confidence: 0.84113336
00:57:47.495 --> 00:57:48.775 shown that, well, can you
NOTE Confidence: 0.84113336
00:57:48.775 --> 00:57:49.275 predict,
NOTE Confidence: 0.94503045
00:57:49.655 --> 00:57:50.455 whether it's a man or
NOTE Confidence: 0.94503045
00:57:50.455 --> 00:57:51.655 a woman? Or or also,
NOTE Confidence: 0.94503045

00:57:51.655 --> 00:57:52.650 can you can you predict
NOTE Confidence: 0.9811446

00:57:53.210 --> 00:57:54.890 the, the race or or
NOTE Confidence: 0.9811446

00:57:54.890 --> 00:57:56.510 other protected subgroups directly
NOTE Confidence: 0.97965175

00:57:56.890 --> 00:57:58.170 from the from the image?
NOTE Confidence: 0.97965175

00:57:58.170 --> 00:57:58.670 Yeah.
NOTE Confidence: 0.99450487

00:58:00.650 --> 00:58:01.849 Can I ask a question?
NOTE Confidence: 0.99450487

00:58:01.849 --> 00:58:02.490 I don't know if you
NOTE Confidence: 0.99450487

00:58:02.490 --> 00:58:03.390 can hear me.
NOTE Confidence: 0.8606285

00:58:04.490 --> 00:58:06.945 Yes. Yes. It's David Klimstra.
NOTE Confidence: 0.9925106

00:58:07.645 --> 00:58:09.345 Great talk, by the way.
NOTE Confidence: 0.8979991

00:58:10.045 --> 00:58:11.005 You know, one of the
NOTE Confidence: 0.8979991

00:58:11.005 --> 00:58:11.325 big,
NOTE Confidence: 0.95748776

00:58:11.805 --> 00:58:13.505 opportunities, I guess, in diagnostic
NOTE Confidence: 0.9983022

00:58:13.885 --> 00:58:14.385 AI
NOTE Confidence: 0.96996194

00:58:14.845 --> 00:58:16.145 is to objectify
NOTE Confidence: 0.99859744

00:58:17.005 --> 00:58:17.505 subjective

NOTE Confidence: 0.7396557
00:58:17.885 --> 00:58:19.345 diagnoses like grading
NOTE Confidence: 0.9973017
00:58:19.920 --> 00:58:20.000 tumors,
NOTE Confidence: 0.97582215
00:58:20.720 --> 00:58:23.119 grading dysplasia, etcetera, etcetera. But
NOTE Confidence: 0.97582215
00:58:23.119 --> 00:58:25.059 of course, because they're subjective,
NOTE Confidence: 0.97582215
00:58:25.200 --> 00:58:26.260 the ground truth
NOTE Confidence: 0.9709586
00:58:26.640 --> 00:58:28.000 in your dataset is also
NOTE Confidence: 0.9709586
00:58:28.000 --> 00:58:29.359 going to be subjective. Do
NOTE Confidence: 0.9709586
00:58:29.359 --> 00:58:31.359 you have any experience trying
NOTE Confidence: 0.9709586
00:58:31.359 --> 00:58:32.420 to resolve that
NOTE Confidence: 0.99636483
00:58:32.880 --> 00:58:33.380 problem?
NOTE Confidence: 0.99664557
00:58:34.365 --> 00:58:34.865 Yeah.
NOTE Confidence: 0.8728895
00:58:36.205 --> 00:58:37.725 Yeah, we have we have
NOTE Confidence: 0.8728895
00:58:37.725 --> 00:58:39.405 looked into it, quite a
NOTE Confidence: 0.8728895
00:58:39.405 --> 00:58:39.905 lot.
NOTE Confidence: 0.9142597
00:58:41.165 --> 00:58:42.765 The the most obvious answer
NOTE Confidence: 0.9142597

00:58:42.765 --> 00:58:44.605 is that is that we
NOTE Confidence: 0.9142597

00:58:44.605 --> 00:58:45.105 are,
NOTE Confidence: 0.9402083

00:58:46.445 --> 00:58:47.725 essentially trying to I mean,
NOTE Confidence: 0.9402083

00:58:47.725 --> 00:58:48.785 I mean, these are
NOTE Confidence: 0.9978296

00:58:49.110 --> 00:58:50.870 continuous biological processes that we
NOTE Confidence: 0.9978296

00:58:50.870 --> 00:58:52.250 have discretized into
NOTE Confidence: 0.9914876

00:58:52.630 --> 00:58:53.450 into these,
NOTE Confidence: 0.9852176

00:58:54.310 --> 00:58:56.010 you know, diagnostic bins.
NOTE Confidence: 0.81341

00:58:56.310 --> 00:58:57.130 And if,
NOTE Confidence: 0.9918606

00:58:58.630 --> 00:58:59.510 if if we want to
NOTE Confidence: 0.9918606

00:58:59.510 --> 00:59:00.650 stick to those diagnostic
NOTE Confidence: 0.8513073

00:59:00.950 --> 00:59:01.450 bins,
NOTE Confidence: 0.9790905

00:59:02.515 --> 00:59:03.575 and there's disparity
NOTE Confidence: 0.9417847

00:59:04.115 --> 00:59:05.475 or subjectivity in the in
NOTE Confidence: 0.9417847

00:59:05.475 --> 00:59:06.675 the diagnosis and there might
NOTE Confidence: 0.9417847

00:59:06.675 --> 00:59:08.195 be some erroneous diagnosis in

NOTE Confidence: 0.9417847
00:59:08.195 --> 00:59:08.695 there,
NOTE Confidence: 0.97819316
00:59:09.315 --> 00:59:10.595 deep learning is a great
NOTE Confidence: 0.97819316
00:59:10.595 --> 00:59:12.115 solution for this because it's,
NOTE Confidence: 0.91352427
00:59:12.995 --> 00:59:15.349 massively robust to label noise.
NOTE Confidence: 0.98451334
00:59:16.049 --> 00:59:17.569 There there are studies in,
NOTE Confidence: 0.98451334
00:59:17.970 --> 00:59:19.089 in machine learning showing that
NOTE Confidence: 0.98451334
00:59:19.089 --> 00:59:19.970 even if you have twenty
NOTE Confidence: 0.98451334
00:59:19.970 --> 00:59:21.730 percent of your labels corrupted,
NOTE Confidence: 0.98451334
00:59:21.730 --> 00:59:23.430 you still get a classifier
NOTE Confidence: 0.98451334
00:59:23.490 --> 00:59:25.029 that's almost perfect because
NOTE Confidence: 0.560795
00:59:25.329 --> 00:59:25.829 the
NOTE Confidence: 0.89479023
00:59:27.135 --> 00:59:27.855 it it can just figure
NOTE Confidence: 0.89479023
00:59:27.855 --> 00:59:28.895 out on its own that
NOTE Confidence: 0.89479023
00:59:28.895 --> 00:59:29.875 what the most common,
NOTE Confidence: 0.9839044
00:59:31.215 --> 00:59:33.135 features are across across these
NOTE Confidence: 0.9839044

00:59:33.135 --> 00:59:34.815 across these images. And if

NOTE Confidence: 0.9839044

00:59:34.815 --> 00:59:36.095 there are outliers, it doesn't

NOTE Confidence: 0.9839044

00:59:36.095 --> 00:59:37.375 adhere to those outliers. It

NOTE Confidence: 0.9839044

00:59:37.375 --> 00:59:38.515 would fit to the data

NOTE Confidence: 0.9927499

00:59:39.230 --> 00:59:40.510 that is most most,

NOTE Confidence: 0.99716556

00:59:40.990 --> 00:59:41.730 most common.

NOTE Confidence: 0.89727795

00:59:42.109 --> 00:59:42.609 But,

NOTE Confidence: 0.9956387

00:59:44.109 --> 00:59:44.609 now,

NOTE Confidence: 0.95163226

00:59:45.230 --> 00:59:46.589 more recently, with all these,

NOTE Confidence: 0.95163226

00:59:46.589 --> 00:59:48.109 like, foundation models and,

NOTE Confidence: 0.94616556

00:59:49.630 --> 00:59:50.770 rich feature extraction,

NOTE Confidence: 0.98052543

00:59:51.230 --> 00:59:52.510 there's an argument to be

NOTE Confidence: 0.98052543

00:59:52.510 --> 00:59:54.030 made that if we can

NOTE Confidence: 0.98052543

00:59:54.030 --> 00:59:55.595 get whole slide level,

NOTE Confidence: 0.9563546

00:59:57.735 --> 00:59:58.635 feature representation

NOTE Confidence: 0.995599

00:59:59.335 --> 01:00:00.695 directly from the from the

NOTE Confidence: 0.995599
01:00:00.695 --> 01:00:02.375 slide, we can begin to
NOTE Confidence: 0.995599
01:00:02.375 --> 01:00:03.755 predict some of these outliers
NOTE Confidence: 0.995599
01:00:03.975 --> 01:00:06.055 directly and immediately without any
NOTE Confidence: 0.995599
01:00:06.055 --> 01:00:07.115 supervised training
NOTE Confidence: 0.91412103
01:00:07.655 --> 01:00:09.355 and have them reevaluated,
NOTE Confidence: 0.94399387
01:00:10.450 --> 01:00:11.490 before you have a have
NOTE Confidence: 0.94399387
01:00:11.490 --> 01:00:12.710 a supervised model,
NOTE Confidence: 0.9348383
01:00:13.090 --> 01:00:14.530 model trained. But we haven't
NOTE Confidence: 0.9348383
01:00:14.530 --> 01:00:15.410 done that yet. That's a
NOTE Confidence: 0.9348383
01:00:15.490 --> 01:00:16.690 that's an idea. That's in
NOTE Confidence: 0.9348383
01:00:16.690 --> 01:00:17.350 the works.
NOTE Confidence: 0.9917523
01:00:18.290 --> 01:00:19.670 Very interesting. Thanks.
NOTE Confidence: 0.9801107
01:00:30.974 --> 01:00:32.815 Yeah. So we we did
NOTE Confidence: 0.9801107
01:00:32.815 --> 01:00:33.955 some work on transplants.
NOTE Confidence: 0.35186514
01:00:47.580 --> 01:00:48.080 So
NOTE Confidence: 0.8924371

01:00:55.385 --> 01:00:56.905 Yeah. Absolutely. We're we're very
NOTE Confidence: 0.8924371

01:00:56.905 --> 01:00:57.944 interested. We we have some
NOTE Confidence: 0.8924371

01:00:57.944 --> 01:00:59.724 ongoing projects including for.
NOTE Confidence: 0.9403819

01:01:01.065 --> 01:01:03.165 We're we're definitely interested. Yeah.
NOTE Confidence: 0.836837

01:01:03.950 --> 01:01:04.450 Yeah.
NOTE Confidence: 0.7844374

01:01:32.170 --> 01:01:33.070 And, you know,
NOTE Confidence: 0.75552344

01:01:33.850 --> 01:01:35.210 and and also a further
NOTE Confidence: 0.75552344

01:01:35.210 --> 01:01:37.130 into that question is, have
NOTE Confidence: 0.75552344

01:01:37.130 --> 01:01:38.810 you guys looked at using
NOTE Confidence: 0.75552344

01:01:38.810 --> 01:01:40.090 any of the AI Google
NOTE Confidence: 0.75552344

01:01:40.090 --> 01:01:40.590 C
NOTE Confidence: 0.66091084

01:01:41.210 --> 01:01:43.470 as a follow-up agent safety,
NOTE Confidence: 0.7295395

01:01:50.025 --> 01:01:51.464 description and correlate to make
NOTE Confidence: 0.7295395

01:01:51.464 --> 01:01:52.924 sure that everything
NOTE Confidence: 0.804032

01:01:53.385 --> 01:01:54.345 that we'll see on the
NOTE Confidence: 0.804032

01:01:54.345 --> 01:01:56.265 description is being represented in

NOTE Confidence: 0.804032
01:01:56.265 --> 01:01:57.085 your slides.
NOTE Confidence: 0.9960669
01:01:57.970 --> 01:01:58.470 Right.
NOTE Confidence: 0.78286546
01:01:58.850 --> 01:02:01.090 That's kinda long. Okay. So
NOTE Confidence: 0.78286546
01:02:01.090 --> 01:02:02.310 so for your,
NOTE Confidence: 0.9454012
01:02:04.290 --> 01:02:05.570 for your second question, no.
NOTE Confidence: 0.9454012
01:02:05.570 --> 01:02:06.530 We we we have not
NOTE Confidence: 0.9454012
01:02:06.530 --> 01:02:07.350 done the
NOTE Confidence: 0.9474716
01:02:10.145 --> 01:02:11.265 the in in in terms
NOTE Confidence: 0.9474716
01:02:11.265 --> 01:02:12.005 of the quality,
NOTE Confidence: 0.9514261
01:02:12.704 --> 01:02:13.984 quality improvement. That is a
NOTE Confidence: 0.9514261
01:02:13.984 --> 01:02:15.365 great great idea, though.
NOTE Confidence: 0.9305765
01:02:15.825 --> 01:02:16.865 I I I think machine
NOTE Confidence: 0.9305765
01:02:16.865 --> 01:02:17.744 learning can be used to
NOTE Confidence: 0.9305765
01:02:17.744 --> 01:02:18.325 do that.
NOTE Confidence: 0.99571216
01:02:19.585 --> 01:02:20.645 For your first question,
NOTE Confidence: 0.8686271

01:02:22.539 --> 01:02:23.039 we
NOTE Confidence: 0.99766624

01:02:23.339 --> 01:02:24.799 are actively working on,
NOTE Confidence: 0.9586287

01:02:25.900 --> 01:02:27.680 you know, improving report synthesis.
NOTE Confidence: 0.9586287

01:02:27.740 --> 01:02:28.700 So far, what we have
NOTE Confidence: 0.9586287

01:02:28.700 --> 01:02:30.400 is that it's able to
NOTE Confidence: 0.94935113

01:02:30.779 --> 01:02:32.779 generate morphologic descriptions from an
NOTE Confidence: 0.94935113

01:02:32.779 --> 01:02:33.680 image and,
NOTE Confidence: 0.9571172

01:02:34.665 --> 01:02:35.645 come up with a diagnosis.
NOTE Confidence: 0.9808469

01:02:36.265 --> 01:02:38.365 More recently, with the context
NOTE Confidence: 0.9808469

01:02:38.505 --> 01:02:40.745 expansion and some some work
NOTE Confidence: 0.9808469

01:02:40.745 --> 01:02:41.405 that we're
NOTE Confidence: 0.935917

01:02:41.785 --> 01:02:43.465 doing with others is is
NOTE Confidence: 0.935917

01:02:43.465 --> 01:02:44.505 is just trying to see
NOTE Confidence: 0.935917

01:02:44.505 --> 01:02:45.785 if within the same context,
NOTE Confidence: 0.935917

01:02:45.785 --> 01:02:46.685 it can ingest
NOTE Confidence: 0.9508386

01:02:47.570 --> 01:02:49.010 the initial image as well

NOTE Confidence: 0.9508386
01:02:49.010 --> 01:02:50.290 as any ancillary test that
NOTE Confidence: 0.9508386
01:02:50.290 --> 01:02:51.430 might need to be ordered
NOTE Confidence: 0.9508386
01:02:51.650 --> 01:02:52.790 leading up to a pathology
NOTE Confidence: 0.9508386
01:02:52.850 --> 01:02:54.450 report and the ability to
NOTE Confidence: 0.9508386
01:02:54.450 --> 01:02:55.490 then go back and for
NOTE Confidence: 0.9508386
01:02:55.490 --> 01:02:57.010 her pathologist to correct anything
NOTE Confidence: 0.9508386
01:02:57.010 --> 01:02:58.290 that needs to be corrected
NOTE Confidence: 0.9508386
01:02:58.290 --> 01:02:59.250 and for the report to
NOTE Confidence: 0.9508386
01:02:59.250 --> 01:03:00.975 update based on on that.
NOTE Confidence: 0.9438024
01:03:01.935 --> 01:03:03.775 But report synthesis in the
NOTE Confidence: 0.9438024
01:03:03.775 --> 01:03:04.815 true sense where it can
NOTE Confidence: 0.9438024
01:03:04.815 --> 01:03:05.695 actually be used is a
NOTE Confidence: 0.9438024
01:03:05.695 --> 01:03:06.595 very hard problem.
NOTE Confidence: 0.9828713
01:03:07.055 --> 01:03:08.415 And it's a hard problem
NOTE Confidence: 0.9828713
01:03:08.415 --> 01:03:09.935 both in terms of having
NOTE Confidence: 0.9828713

01:03:09.935 --> 01:03:10.835 enough data
NOTE Confidence: 0.9956496

01:03:11.215 --> 01:03:12.735 to because you can't just
NOTE Confidence: 0.9956496

01:03:12.735 --> 01:03:13.235 use
NOTE Confidence: 0.9651653

01:03:14.730 --> 01:03:16.410 the existing pathology images and
NOTE Confidence: 0.9651653

01:03:16.410 --> 01:03:17.770 reports to do this. You
NOTE Confidence: 0.9651653

01:03:17.770 --> 01:03:20.030 need fine grained morphologic description
NOTE Confidence: 0.9806369

01:03:20.730 --> 01:03:21.230 at
NOTE Confidence: 0.78040427

01:03:21.690 --> 01:03:22.430 at at regions,
NOTE Confidence: 0.9521383

01:03:23.850 --> 01:03:24.350 and,
NOTE Confidence: 0.9300672

01:03:25.770 --> 01:03:27.210 and and with evaluation. Right?
NOTE Confidence: 0.9300672

01:03:27.210 --> 01:03:28.170 So it's it's very difficult
NOTE Confidence: 0.9300672

01:03:28.170 --> 01:03:30.326 to evaluate as well. Yeah.
NOTE Confidence: 0.81525415

01:03:30.946 --> 01:03:31.666 I think we have five
NOTE Confidence: 0.81525415

01:03:31.666 --> 01:03:32.546 minutes over. Thank you for
NOTE Confidence: 0.81525415

01:03:32.546 --> 01:03:33.346 a great talk. Yeah. You
NOTE Confidence: 0.81525415

01:03:33.346 --> 01:03:34.006 can see.