So, good afternoon.
First of all, we welcome students from the other class.
So, we are friendly classes.
So, this is the last one.
We are in different sessions, but pretty much we follow very similar schedule.
The last one, we discuss with your instructor and we think machine learning is an important topic and we may introduce a little bit general background, some new trend.
So, students have a widened vision about medical imaging.
So, this is combined class.
This is the experiment.
We are not sure how much water would be the best way, but we do believe machine learning is major direction.
And that's a wonderful thing to know little bit.
So, let me try to explain machine learning in the context of medical imaging.
And there will be three parts in this lecture.
First part is the background in an earlier class and we explain the network.
In this particular session, we explain both electric network and also a neurological neural network and artificial neural network.
So, we first review a few slides and those students who are not in my class, so at least you get some general feeling about artificial neural network.
So, normally I just say neural network that really means artificial neural network or computer versus computer model of neural network.
And then I gave some exemplary applications outside medical imaging field.
So, this is the first part.
The second part we really zoom in, medical imaging.
So, water is the relevance of machine learning to medical imaging.
So, I will present my general perspective and some ladies to result.
And finally, I talk about the future of medical imaging.
And last two parts pretty much reflect what we have been doing at RPI.
So, we are leading group in this emerging area.
So, hope students will share our excitement.
So, this part I will finish with a summary slide.
So, if you remember those three summary slides, those three key points.
So, you will bring I think good information home.
So, first background part, we talk about neural network.
So, water is neural network.
We start with biological neural.
So, we have intelligence.
So, based on biological neural system.
So, we go back to the very fundamental level.
So, you have a single neural.
So, single neural is a imaging structure.
So, you have biological stimulation.
So, come out as input into a single neural.
This neural is small separate processing unit.
If you like using a small biological computer, very complicated.
And even just single neural.
But very roughly, you think neural networks take input, all biological stimulation and input into cell body, neural cell body for the summer.
So, so called dendrites.
So, this is the multiple wires, branches.
So, the information will come into the processing unit.
And the neural will synthesize all the information.
And depending on the patterns, strands of stimulation, the neural will decide if we will get excited or stay positive.
So, get on or off.
The signal on will introduce so called electrical potential, which is electrical pulse will be sent down the through the fiber and biological term is a x-on.
So, the signal will come down the way and into other neurons.
And the interface is synapses.
So, very complicated thing.
So, this is single unit.
Many such neurons connect together.

Then you can do amazing things like you touch some very hard surface immediately.
Just the reflex back.
Or you see things, you recognize the student coming or you are taking notes.
So, these high-level activities all trace back to individual neurons and they are interconnection.
So, this is a biological thing.
That's only biological thing I would mention in this class.
Then the mathematical model of this biological neuron becomes artificial neuron.
So, mathematically, you view this as two parts.
It's interesting thing.
It's two parts.
So, this part means if you imagine a line, just draw a dotted line.
So, separate neuron into two parts.
The life part, this part is not seen, but a linear form of information processing.
So, you have say three important x, y, z.
Then three important will be weighted by a x, a y, a z, three coefficient of code weight.
Okay.
Then the identity together.
And this is the sum together.
So, this is the a x times x, a y times y, a z times z.
So, this is in linear form.
And also in the foundational part, we know this is in a form of inner product.
So, inner product is a linear form.
You do linear transformation.
So, we learned linear system theory.
So, we know a lot about linear system and convolution for analysis.
So, all linear system parts.
So, you see inside all the information through inner product, which is a linear operator.
As a single number, this single number then subject to a 90-year transform.
I will talk more about the 90-year transform.
So, depends on magnitude of inner product results.
The output can be high or low.
So, this is 90-year transform mission.
Some seem like thresholding.
If the value to small, ignore it.
If too high, this report, I get excited.
So, this is something like this.
The small thing is a combination of linear and 90-year systems.
So, you have two parts putting together to map what we learned from biological neurons.
So, this is a simple unit.
And the neuron also called many years ago called perceptron.
So, you senseless same thing.
So, you see multiple input values.
Weighted by W1, W2, W1.
If you particularly make one weighting factor equal to one, then the weighted results will be a bias.
So, you have a bias.
You can just treat it as linear form.
But in two parts, one part is inner product.
The other part is a constant off side.
That's a matter.
This is like a linear equation.
You have intercept, this A, Y equal to KX plus B.
B is off side.
So, for neuron, you could have this off side.
And this off side could be combined into inner product form.
So, this is a simple linear processing part.
Then we say you need a nonlinear activation function.
There are many kinds of nonlinear functions serving as nonlinear transfer

machines for artificial neuron.
There are multiple things and some listed here.
And I just copied a table from the internet.
So, activation function.
Multiple functions.
And most frequently, sigmoid or self-function.
So, you basically have something like this.
So, if the input is very negative or small, it gives you zero.
So, if very high, it will saturate.
And just give you one.
So, if you could see this carefully, you got something like this.
Okay, this is sigmoid or this special tangent function.
We will do the same trick.
This shape of the nonlinear function makes most signs.
And a very popular nonlinear activation function is so called this rectified linear unit or rec linear unit.
So, basically, you have all straight lines, straight at zero.
All of a sudden becomes a linear function.
The slope can be changed.
So, this is the rec linear unit.
And nowadays, we do medical imaging research.
We use this rec linear unit most of the time.
So, this is very cool selection and very good numerical property.
And single neuron is important, but not most useful.
And in practical applications, oftentimes, we need link multiple neurons into a network.
Multi layer feed forward network.
So, it is very typical.
So, you have input layer.
So, input value is say x, y, z, all you call x1, x2, this is so on.
And you wait by different waiting factors.
So, this is is input, weighted by w, yj is goes to gs unit in the next layer.
So, for each layer, you have input.
So, you can have fully connected neural network shown like this.
And this is next layer.
Next layer can be connected to the next layer.
So, you can have multiple layers.
And usually, the neural network has been around for several decades.
And usually, you only have one layer or two.
But nowadays, talking about deep neural network.
So, you have many, many layers.
Then, final layer is output layer.
Input layer take value, you measure it.
So, you know, and output, you observe.
So, you get output value recorded here.
And the middle layers are not directly observable.
Therefore, they are called hidden layers.
So, you have input layer, you think you have a vector x.
You have output layer, then you have g.
Then, you have middle layers, you call y1, y2, or vector y.
You have multiple layers.
So, this kind of structure is very popular nowadays.
But you can imagine, you can connect in different ways like graph, like three many things.
But we do not have time to talk all kinds of possibilities.
I will mention little bit.
But we focus on this kind of multi layer feed forward neural network.
And later on, we will introduce some sort of path and so on.
That's the mindset very briefly.
Okay.
So, artificial neural network.
So, the problem is what should be the neural network architecture, how many layers, how many shells, and also what should be these weighting factors.
All these are unknown.

And the literature indicator, if you have good selection of number of layers, number of shells, neurons, player, and if you have good way to adjust the parameters, you can do amazing things.
As amazing as you can drive automatically without human intervention.
So, this is really amazing thing.
But fundamental idea, how you select number of layers and neurons, I would say at the moment, this is a big open question.
We do not have any governing theory to say you have to select, say 10 layers or 12, what so I will, we do not know exactly.
And once we try on the iron, we give you a neural network, just say this may work, just try.
Then you still have a question, how I set all these weighting factors, W's, many W's, maybe milling W's.
And these values initially, you just randomize, just put random numbers.
For example, you normalize your input, and you just select all these W's, random number between 0 and 1 or from minus 1 to positive 1, you can do many ways.
Once you guide the parameter, initialize in the way, just mention, and you got input, and you have the output.
Output could be, just give you simple example, could be like simple label, input is vectorized image.
So you take image, image, there is a cut in the image, there is a dog in the image, you have milling images, containing, user, cut or dog, you put image here, and the output, you x back, you put cut picture, the network will be, you take it, otherwise it takes dog, so this is simple, cut dog classification.
So you have big data side, many, many images already labeled as cut or dog, or maybe you think could be more complicated, say digits, 0 to 9.
So you could have multiple output depending on application, the simple things like this binary decision, but what I work, you have input, and because you have big data side, because you already labeled your big data side, so give input, you know what the output should be, and because we randomize, randomly initialize the network, the chance, you give output, this is based on you randomly, randomly side parameters, you give input, say cut, the output is 1, that means a 0 means dog, so normally you output from your initial network, wouldn't mind your ideal output, or your label, so there will be aero introduce, so you add all these aero together, based on training data side, say 100 images, I put 100 images, CE Aero network give me 0 or 1 mean title dog, and sometimes it tells me correct answer, many times just wrong, so I measure all the aero's using least square for example, this is not only way, but this is good way, you see, given network, you can define an objective function, that's a aero, aero is big means you know the network should be like this, five layers, easy layers, say certain number of neurons, but because the parameters were not correctly set, so the total aero is still significant, then what you will do, this is based on the aero, you change your parameters, you see the derivative will change the parameter, you do you do gradient descent search, so just the perturbed the current parameter vector could be many many weight parameters in the vector, you perturbate little bit, so you want to see how the total aero will change, you perturb certain direction, the aero will increase little bit, the other direction the total aero will reduce, and you just do numerical optimization, you find steepest descent direction, then you change your parameter little bit, so this is a direction along this direction, this is the parameter vector, many parameters, so this is really a greeting operator, so along the greeting, you just proceed a small step, define by learning with eta, usually eta between 0 and 1, you make a small step, then you modify your current parameter along that steepest descent direction, so this is, this will give you a little bit smaller total aero, so the current parameter is modified, according to greeting direction, then you have updated parameter, okay, this parameter will by definition will reduce the objective function little bit, so this is a graduate requirement process, so you do once, you get little bit less value of total aero, you do it again, again, again, what will happen, so roughly, this is your starting point, okay, this is your starting point, okay, along the greeting direction, so you do small step, refinement, so you get little bit less total aero, you keep doing so, and the graduate will reduce the minimum, okay, this is certainly a simple case, only two aero, and for big neural network, deep neural network, you have many, many more parameters, but visualization is the same,

something like this, so this is the, let you know how neural network and the guide good performance, and again, I mentioned that we do not know from fundamental governing theory, how many layers of all the kind of connection architecture, we do not know, and we know we can do the search, and you do the search, you still have problem, so you could end up with minimum point, this minimum point may give you local minimum, may not be global minimum, there could be many local minimum point, but the purpose is find global minimum, how can you guarantee your numerical search, the heuristic searching process, as I described, will give you global minimum, we do not have theoretical guarantee, so this is an active area, many things we do not know, but good news is that surprisingly, this deep neural network, this heuristic search, and oftentimes, almost always give us very good results, and why we keep having good results, we do not know exactly, this is again an active search area, we have some semi-contactive understanding, we have some inside, so internet, inter-pertability, and overall convergence, and why we have practical performance, what's reliability, we do not know, but this will lead to next big breakthrough we believe, and somehow the neural network like black boxes, so people feel do we really want to use it, we do not know what's going on really, but it's performance is really good, like people have will and do the statistical model, you can predict, stock market performance to a certain degree, and you build machine learning network, you can do better, so we sometimes will use it, so one way, classic way, you know fundamental assumption, and you understand the old governing theory, there are other one neural network, you do not know much, but it do better job, it will do better job for you, so you know the answer, so this is current situation, so let me just recast in different words, so traditionally you have bands of data point, these right point, and somehow you know governing theory, this is quadratic function, then you try to estimate the parameter, you do quadratic fitting, so something like this, this is a classic problem, I would say neural network is a sophisticated fitting measure, and the data much bigger, dimensionality much larger, and you still want to fail, and you do not know what will be the boundary is not quadratic form, not probably is not say single solid form, it's a rather complicated, and you do not know analytical expression, but for many problems, universally we can use neural network, this is a simple neural network, do digit recognition, so you use all these things to put together, so if you do high order problem, you either do high order approximation or do quadratic fitting, you know the functional form, but here the functional form is really combination of many neurons, and each neuron is a combination of linear and nonlinear unit, so these things you can just do layer by layer, so initially something like wave light analysis from lower level, then you see this as the information to higher level, it just work fine, so this is just a way to think of machine learning as a sophisticated fitting measure, and the labels, these are data points, but for machine learning we can label, so this is three, this is dog or water cell network, and many human level activities can be performed using neural network, and there are some very fundamental x-plorizon, basically it says, with neural network, and you can represent any general functions, so this is the one block to show you graphically how neural network can solve any problem, and this is practically any problem, you can just arbitrarily draw functional form, again this is one dimensional situation, then you can fit the functional form with so-called reclinier unit, this is the reclinier unit, it's just street horizontal line, and all of sudden you jump, so something like that, water show you, so this is all zero, then you just get excited, so this is the threshold, and this value functional value is one, then keep going, so if you add multiple reclinier unit, which is different slope, different off-side, you could put it here or here or even flip it around, you use simple functional form like this, you can do good fitting to this general continuous curve, so this is just show you example, if you want to look at more detail, you can check the block, it will show you one step after another, so this is to give you an idea, and with neural network we have some theoretical inside, any function, any, this means usually practical function, can be real fitted by neural network, so any processing or intelligent activity, we think that is a form of computation, then it can be performed by a multivariable general function, then the function can be implemented by neural network, so the neural network is to be good pros, then the problem is how you design the neural network, how you optimize its parameters, and I already explained a little bit, so this is just very fundamental knowledge little bit, so you know, and the human intelligence curve works more than simple things as I just showed you, for example, recognition, the human vision, and it's a good part of human intelligence, and the human vision arguably gives you primary source of information, so you are asked to recognize this is chair, so you just think how computer machine will do that, and you can keep computer, this is chair, computer remember, you can keep changing lightning condition and angle, so on, then you see another chair, okay, you see the problem, I think, more challenging, and this is even more challenging, these steel chairs, so you know the intelligence, computer reason is not easy thing, for many years no measure progress, until recently, for past several years, and the image night is a huge database, and multiple million images, very labels, and the computer scientists organize context, try to write human performance level, be challenged by computer vision algorithm, you see very recently, so 2000, just say, it's 2011, so see the IREAT quite high, and 2015, not long ago, so the computer performance beat human, using deep neural network pros, so this is a simple example, showing you impressive progress achieved by machine learning method, and also automatic driving pretty much image recognition problem, okay, so you see this is on the way, and people believe in next few years, maybe five years, maybe ten years, and we will have fully automatic cars, define the automation by several levels, and I myself very much want to have one because I'm not good driver, okay, so not only image recognition, also game planning, and computer game program, alpha, gold, zero, and the kind, it's human player, so this is just a few examples, there are many other examples, recently, good paper, good paper about skin cancer recognition, and right now, you take iPhone picture of your skin problem, and the diagnosis is generally better than trained physician, so many exciting stories, people got so excited, and the idea, so just summarized here, and we need to use deep neural network, this is just feed forward multi layer neural network, you should not stay waste too, you need one or two, you need to use maybe 20, maybe 100, 200, so deep neural network is a way to go, versus silo neural network, deep neural network need to be trained, so the parameters should be optimized, how you train neural network, you need big data, you need good optimization algorithm, so, and after you train the neural network, when the data size is large, it's small, so traditional like statistical method, all the silo neural network measure, or deep neural network measure, if data size is small, they more or less the same, but when data size goes exponentially larger, deep neural network is clearly the way to go, artificial intelligence is a big domain, then machine learning is one part, major part, then deep learning is mainstream of machine learning, so this is really enabled by big data, super computer, and the computational scale, and so on, so specifically, so as time goes by, say from 2010, until say a few years ago, you see the performance, guiding better, in terms of top 5, I wrote, so what do you mean by top 5, I wrote, so I define here, correct answer is among the top 5, top 5, guys is made by network, we think the network is doing reasonable job, okay, so the I wrote guiding smaller and smaller, until this point we beat human, so this context stopped, because we already beat human, but you see on the other hand, the number of layers of neural network, getting more and more a few years ago, you got over 100 layers, and several years passed, and you have more examples, this is very, very hard, I would say probably how this I wrote of the search, so the deep neural network, the complexity of the system, and the guiding increase, so you have stronger representation capability, and you can train and waste big data, and the previously training data side is small, so nowadays the data side getting larger and larger, so you see the trend, and yesterday I had this class, with our biotank building and the city micro city facility manager, so we started archiving all the micro cities, so we have multiple groups doing small animals, that is, this is going to be a nice, and the year we began about 100 miles, and the raw data images about 10 gigabytes, so we want to save all of them with certain reliability, so we request a data server about 200 terabytes, so we have all these data, big data will be used to train neural network, so we can do better, image reconstruction, and image analysis, and you have big data, you have complicated network, and you also need high performance platform, so computing power also increases exponentially, so this is just in parallel to data size increment, so computing power becomes more and more impressive, and traditionally you use CPU, so now you use graphical processing unit, or even tenser processing unit, and these new things, graphical processing unit, or tenser processing unit, it's not as good as traditional CPU or computer, if you just do sequential computation, just say I calculate pie, so in thousands, data, precision, after that's more, and then CPU will just do good job, okay, the GPU or tenser processing unit, graphical processing unit, really good for power line processing, so this is very suitable for machine learning, and you can power line algorithm, and result GPU or tenser processing unit, and you do not have efficiency in training neural network, and in my life we have multiple GPUs and TPUs, and we train a network, it may take hours, days, up to one week, and once we finish training, the network will do the job very efficiently, so these are some terminologies you need to know, and when you do machine learning, and most popular language is Python nowadays, there are several competing languages, even MATLAB is a good language, but so far Python is very much popular, I wouldn't say dominating, tenser flow supported by Google is platform, so you can you can develop machine learning algorithms using Python in the on the platform of tenser flow, and because the platform was dedicated to machine learning, so this starting point is in the data high, so if you are interested, you spend a few weeks or months, you should be able to do decent job, like Google example, those of you really interested in this, you want to guide your hand, hand on things, my suggestion, you just read this blog, okay, and you try to learn tenser flow and deep learning and the Python, so you need to learn Python, Python pretty much like C, like any language, and you read a few days, you know Python, okay, and you read a few days, you know tenser flow, you follow blog, you see some examples, you read your honest, then example, you just modify those examples, it will run for you, so you have many internet resources to do, to do say digit recognition, so this I'm next, data side, just give you big data, all the numbers, images, I used by why I couldn't remember, so already labeled, you use these data side to train your neural network, different kind of neural network, we have different performance, and many heuristic tricks, numerical techniques, and put the guide, I think within one day, when you can just reproduce these high-level performance, it will recognize digit, that's very amazing, so on one hand, this is a fancy topic, on the other hand, the threshold is not very high, so this is something, and we are doing next part, I talk about machine learning for medical imaging, so in our field, and we analyze the situation, so machine learning is hard, we want to use machine learning to do medical imaging, we should not wait for computer scientists, or other people to do the machine learning based on medical imaging, we as imaging researchers, should learn machine learning to medical imaging, why is that, and we think, the way we learn machine learning is relatively easier than other people, than medical imaging, you need to know, I'm not, I say, very confusing, right, you have a lot of MRF, X3 polychromatic beam, heartening, and forest, Lycoceros, combing beam reconstruction, divergent beam geometry, how you handle data transmission, let me take several years for them to reach the level, to do research, and this is, this is take us at most few months, so we can use machine learning to do research, so this is not symmetric situation, so we just say, let's do it, and next part, I will mention deeper reconstruction, how machine learning becomes relevant to medical imaging, or medical tomographic reconstruction, so before, go to next part, let me give you first, take a home point, so summary one, coding versus learning, and you know how do you do coding or programming, you have data, you divide it for program, okay, the computer, tutorial program, and the take data is input, then the program process data, give you output, this is a classic view, so now we are using machine learning, so we say, the program, the model, is not explicit to you, rather you like the computer, learn from big data, so you have data, the data already labeled, labeled is what you, you output data, you expect the system of network will, will report, so you have big data, and then you, you let your network be trained by the big data, then the outcome will be a network, trained network, or a computer program, once you train the neural network, that is a program, so you can, new data comes in, the program will process the data, we will say this is cat, this is dog, this is CT images, and we can just produce very smart results, so this is kind of, part time change, so you need change, you mind the from coding to learning, you like it or not, this is happening nowadays, okay, first part, the second part is not general knowledge about neural network, now we talk about reconstruction, so two or three years ago I wrote a positive article, positive article on different imaging, so what's the difference, and what's the new twist, we are trying to do, so you can search this, search this, and you just input the title, google it, you will find full paper, and if you like, you just read, and just let me know, if you have any comments, so publicly known exercises of machine learning can be very roughly summarized as this, so you take an image, so cut dog, then neural network got trained, and this neural network is so powerful, it can correctly recognize, remember image, night, complex, all about image recognition, you starting point, you have images, so you input images into a network, output is feature, or label, this is a kite, or just the big eye kite, or just the long arm dog, what's the other way, so input is a image, output is feature, a feature or feature vector, so image to feature, so pretty much all the hottest stories, public, new, public, media, my exposure, about images to features, pretty much, but tomographic imaging wise, and we really produce images as our end product, so our output is images, what's the input, input is feature, so it's just the opposite first size, so from feature to image, the feature can be Fourier components, harmonica components, can be line integral, can be reflected, icons, scattered, back scattered signals, in ultra sound imaging mode, can be emitted, gamma ray signals, so many things, or optical diffuse signals, so this is not the same, and the computer vision has been publicly known, but computer vision only see superficial image, this is like take picture, kite or dog, you can now see through, tomographic imaging is a step deeper, we want to see what's inside, so I say, we're learning one of the next frontiers, is tomographic imaging, so this is what we are trying to argue, and based on this perspective, article, IEEE, Trans-Sexism, medical imaging, organized special issue, so we just try to collect good papers from research group, we are closing this case, so I'm supposed to finish an editor, editor, real draft, within this mind, all about tomographic imaging, so this is x3, this is ultra sound MRI, so this OCT, so any raw data, put into neural network, it will guide images, you can use it for individual imaging modality, or you can use for multi-modality, so many ways to utilize neural network for tomographic reconstructing, so more generally, this is the big picture as I perceive and reported in the perspective article, I call it smart medicine, so you see, today's, and you have images as you starting point, so you can do something like computer vision, so you can just guide images, neural network to segment images, to tell you where is the tumor, where is the liver, where is the prostate cancer, what is the so this is somehow like image processing, or post processing, so somehow, over-life is computer vision, so this link is go from image to features, so TMI, I mean, TransSexism, medical imaging, organized special issue, all about image processing, so what I propose is to go this way, so you have data acquisition, you have raw data measurement, synograms, case-based signals, water-soluble, so you just use neural network, and to do good job, so from these features, you can guide tomographic images, so this is a different utility, as water is mentioned earlier, so this is utility tool, so from data to image, so if you can go from data to image, and from image to feature, you just combine the two precise together, you streamline the two steps, then you see it is certainly possible to go directly from data to features, so this is this link, once you have images, you have features in future, near future or remote future, the intervention like robotic surgery, should be image guide data, so the image should be analyzed into features, guide the surgical tools, so you have utility three, so the features will guide therapeutic steps, like robotic surgery, that's very critical, all some even you take some drugs, you want to monitor, if you have good response and so on, you still do need to do some sensing imaging, so this is certain utility, so from features to action, certainly you can imagine, and you do not need to go this round loop, to image, to feature, to action, and maybe from data, you can use black box, called neural network, just do some training, then the neural network will target what's next step, black field, what's next step, you should take, so this is overall picture about smart medicine in the framework of machine learning, and let's just take about 10 minutes, then I will finish the rest of the lecture, okay.

Okay, let's continue, and yesterday I tried to upload this slide, better memories, it's got overflow, and we will fix the problem, so you should able to see the slides, if you want to review, so I just explained this overall picture, so the next step will focus on this link, this green link, from data to image, I will mention little bit, image processing, this one, but mainly I focus on this link, and later I will say how we go from data to feature directly, so from tomographic data, like pat data, or just like x3 projection, or xbin iql data, so you have row data, rows data cannot be directly interpreted, and you cannot visualize data in a diagnostically meaningful way, you look at the signal ramp, you do not know where it tumble, you need reconstruct images, then radiologists can read images, so from row data to images, form say, i-ray code image formation, or image reconstructing, and oftentimes we heard, they say, in the tomographic reconstructing i-ray, there are two kinds of methods, one is an electric method, or project, so like what we learned, filter the back projection, it's a formula, just one line, so this is a closed form solution, so this is an analytic process, you give data, we plug in data into the formula, you do bunch of computation, you got image, filter the back projection is a good example, and certainly, an analytic approach, and oftentimes, it's not as simple as filter the back projection, as we learned, and I mentioned in combi-mjome, you do helical scanning, the mesimatic is very complicated, you need to understand the high level mesimatic to follow the direction, but after all the direction, the formula is still one line, this is one line double integral, and we can interpret it as small, the ones the filter the back projection, so there are many ways you have a formula, so an analytic approach, the second approach is called a theoretical approach, when we learn nuclear tomography, and we are exploring how you make initial guides, and based on initial guides, you check this frequency between your major data, and you see the size of the data, based on your intermediate or current estimated image, and you see this frequency, based on this frequency, you go back to refine pixel values, your update, your current image, or your estimated image, to be better, so keep doing it recent, then the image gets better and better, and the other reason can be guided by an objective function in terms of least square, so just say this is measurement, and this is what I would predict, based on the current image, the current image is not perfect, so you do projection based on current image, you projected the same size of the projection, and the water you really measure the projection, they do not agree, so based on the difference, you go back to to modify the current image, so you can see less distance frequency, and also you can use maximum likelihood as your objective function, you can keep going, so many ways to guide the iterative precise, well as you, so that the tomographic image can be refined step by step towards the final solution, so a vast literature about iterative algorithm, we need to know what should be the objective function, how we help iterative precise to convert to the true solution, water-aif, the data is not complete, and water-aif we have strong noise or bias, so many things, you can use information theory nowadays, you can use so-called comprised sensing, or loading sensing, or manual for the learning, all these beyond the scope, so iterative reconstructions, algorithms can be still representative, or one line or two, so just keep doing the same thing, again, again, again, so after each cycle, you get a little better image, so previously we think only these two processes, analytic or iterative, iterative is time consuming, but usually works better when data not perfect, like a load of CT, so we use iterative algorithm, we can deliver decent image quality, but at cost of computational overhead, with machine learning, now we say the third category of algorithm emerged, called learning process, so you use neural network, and previously in hospital clinic, you do CT scan, and you guide the diagnosis results, then your data, your image, this is the data category, this thing we do not use anymore, and as I mentioned earlier, now we try to archive micro CT data and images, and many hospitals they now try to save medical CT, MRI images, and the diagnosis, electronic report, so all these are very valuable, are gold mine, so the data server is very popular, so you have all these data, so all these data represent very extensive individualized applications, specific knowledge, and with that knowledge, we can do better job in the framework of machine learning, and I mentioned neurons as I explained can be linked as so-called feed multi layer feed forward network, so this is deep feed forward network, this is only thing, I mentioned so far, there are certainly many other possibilities, you can even make a neuron has certain memory, so the neuron reacts not only depends on current input, also depends on previous status, you can kind of have memory, many possibilities, and as widely possible as you can imagine, you can link to any way you want, not all the connections, not all possible neural network actually textures are equally

useful, and we do not have common theory, but we have engineering and scientific inside into neural network, and we want to have some guidance based on what we learned before, like we learned wave analysis, so we know the multi scale analysis, first you analyze small features, based on small features you see size, intermediate level features, and goes dive dive, this is wave light transform, wave light transform is linear system, so everything is linear, but neural network I think is the one form of wave light transform, if we form a network, this is the multi layer feed forward network, and in light of wave light analysis, you think first layer is just a guide small features output, but then the next layer you can synthesize the local fine features, I extracted using the first layer, so you do this dive dive, and indeed this is good interpretation, if you check internet those phase recognition tasks, if you just visualize what is the output of an arbitrary layer, you see lower level means the layers closer to input, and give you local feature, like first layer give you like edge corner point, then step by step, at certain the lines the layer you start seeing eyes, nose, and finally you just match to individual phases, so this multiple scale analysis is good inside, then we can understand why you have multi layer feed forward neural network, because you want to perform multi scale analysis, just the neural network is more powerful, it's not purely linear, linear is very limited, as we learned linear system theory reminds chaotic and nonlinear system, so now we see the neural network is elegant way to combine linear processing and nonlinear processing at a fundamental neurological cellular level, so even simple unit you have linear and nonlinear thing combined together, so you have the power from nonlinear processing with neural network, and also we learn linear system, we know convolution operation is very good, we learn the equality, we know convolution can be used to detect features and do match the optimal filtering, so on, so the neural network, the inner product part can be casted as convolution, and that way you do not have many parameters to fit, because convolution is processing based on safety and vulnerability, so you have the parameters, good here also good there, so the kernel is not changing, because the system is not not specifically or temporarily wiring, so you reduce number of unknowns, and also reflect any benefit, we usually have with convolutional operation, so this is kind of related to what we learned about linear system and also in control theory, we know the feedback is important, so in neural network we can have some sort of cut link, so to have the benefit from feedback, so you can make system more stable, and also more advanced topics like game theory and the weekend design, so called other virtual neural network, so you can make neural network, two neural network, they play together, so by neural network try to say generate some images like dog, the other neural network just try to tell difference between real dog and dog you generated, so they compete each other, you have a this word other words really lost, so we just like two neural network compete, then when the system can work, can work, then the generator can generate a realistic dog image, and the disk generator will be very soft, Thailand difference, so through this synergistic competition you end up with good neural network, and now the neural network as I mentioned take input and weight, or in real values, so you can imagine you can make input and weight or complex numbers, and you still remember the beautiful formula, your formula, so you can have complex build into neural network, and amplitude and phase comes naturally for medical imaging, like x3 imaging, if you do face contrast x3, you have amplitude and phase, you do electron magnetic wave, and you have phase and amplitude, so complex numbers are good way to represent the physical reality, so this is to say you have many neural networks, a lot of possibilities, the key is learning capability, how do you configure your neural network depends on your application, and you inside, before we have universal theory I thought we will have success theory in the near future, you can based on what you learned previously, and try to build the neural network, keep trying, so through experiment you will likely get very good solution, and the machine learning based on medical imaging is not only a academic interest, and industrial firms, major
medical imaging vendors like general electric, Siemens, Philippe, Toshiba, United and American in China, the neural software in China, many medical imaging companies are all interested in this direction, and as our neighbor, general electric global research center, very close to us, and the world leading expertise, and engineers, we are in collaboration, so from January this year we

formed a strategic lantern pattern shape to develop deep neural network probes for CT, medical CT, we are working actively on machine learning now, so a lot of development I cannot say a lot, this industrial research is different from a academic expert reason, so you need to hold certain things until it becomes publicly available, so now let me show you some example, what we did, so this is a paper, we say how we can use neural network to remove CT, the image noise, when you use the load dose CT scan, and you save the release in dose, but the data becomes very noisy, if you just use filter to backup projection, or even use iterative reconstruction, the image still looks noisy, and there are some artifacts, so we build a neural network, and we call it, it's a residual encoder, decoder, convolutional neural network, which is a feed forward network, so about 10 layers, so you go layer by layer, just go this way, so you do convolution, it's not just fully connected, it's arbitrary waiting factor, you really implement convolutional operation, just like what we learned in the linear system part, the convolution thing, the same thing, then the convolution sub-sector, this glyclineal unit, I mentioned this is output into next layer, so just feed forward, nothing more than what I told you, but the different thing is that you have the unique feature, these links, so it's not just fully flow, layer by layer, and some output directly feed into some later layer, just jump to layer here, so the inside is something like feedback link, so if you do jump like this, so what learned is the difference, so you focus on changes, so this is something you can do at multiple-regilient level, so this kind of more complicated system, so you try to do comparison with the encoder part, then you expand the feature, through the decoder part, when you access the feature, the system is learning, try to keep mostly important feature, so those noise artifacts will be surprised, then you expand back, guided by big data, so all these model parameters will be trained to give you good results, so this is just a visual example, you see the current noise images, and you do mostly learning pretty much clean up the noise, this is not perfect yet, and we have a collaborator, radiologist from half-form medical school, and he pointed, it's still not, like features that are going to take information, preserve the pretty much, but their critique are too, one is that the series, tiny spot, the medical doctor, they think here you look carefully, the small black dot, it's very dark, but after machine learning, it becomes gray, so he doesn't like that part, also he think these artifacts, something like blocky things, they could be improved, but overall gave very positive feedback, and also the over-smoozing shown here would not be preferred by some radiologists, so we did some further improvement, and try to keep features, and maintain natural looking, so this is some latest results, and we submitted to IEEE Trans-Sexymetrical Imaging, so we do so-called contracting paths, and we do not call it conveying paths, so this guy features brands over, and we try to train, this is densely connected night, so this is the filter, the back projection, very noisy result, this is the other common from the master reported in the earlier people, so this is a densely resolved, we obtain, so the image looks more natural, still these key features retain very well, so these two examples, still doing image positive processing, then the positive processing is not the major point, we really want to do data to image transformation, like the 2D Synogram, and we want to make first the convolutional layer to guide the filter layer, so this is filter layer, so this is very natural because filter the back projection, talking about basically invariant filtering, so we just might the formula into a neural network layer, this is convolutional layer shown here, after this layer you just do, do sparsely connected layer to form image, and this layer carries bound to back projection, back projection, so filter the point, the value will be similar the back over field view, so this only along this line, so data is a sparsely connected layer, so in reference to filter the back projection formula, you can immediately form this layer, but why bother you call this FBP neural network and several layers, not just do filter the back projection, the benefit doing this is not, once you form the neural network, so all these weight can be trained, not necessarily today the same as specified by your formula, because of this additional training capability, the imaging performance is expected to be better, and we ask a deeper question, why you want to refine filter the back projection formula, and this is actually very good question, filter the back projection formula, is based on mathematical model called line integral, okay, the line integral model is only a approximation, and we know the CT data formation, and that is really described by double integral, why integral is line integral, the other is integral based respect to x3 energy, because the medical CT tube I made is polychromatic x3, so the line integral model is only approximate, and not only that you have focus, you have detector response, many factors, you cannot possibly all included in your mathematical model, any mathematical model is only proclamation, but with machine learning, you do and to end training, you know what's a good standard, like you scan phantom, you know what should be, so any of those I-roar model mesimites, what is the I-roar, will be automatically learned in the neural network, that's why learning method is a better method than traditional iterative reconstruction method, so you try to utilize extensive knowledge with this highly flexible network, you train, suppose you can find good parameters, so this is just the idea, you do an analytic neural, analytic algorithm based neural network, and these people just got accepted, and this is a neural network based on iterative reconstruction algorithm, so iterative algorithm, you have a current image, through certain mathematical processing, you got to refine the image, then this refine the image is defined as a current image, then through same process, you got further refine the image, you keep doing it, okay, use staff just like a forward analytic formula, you keep doing it, and usually you need to do hundreds of thousands of times to get good image, and here we just take ten iterations, and the easy iterations is just like an analytic reconstruction, just a closed formula, so we just take ten layers, and each layer we might into two parts, one part is the iterative algorithm, you have a formula, right, how you do the refinement from a current image to you refine the image, so this is the formula, as an example, you do this, this is the current image, you do so, you got to refine the image, but because aside any model based on which you derive your analytic solution, or you derive your iterative reconstructing formula, it's not exactly, so in parallel to this iterative formula, and we have a path which is formed by multi-layer neural network, so with this layer, and we can take a care of any model mismatch, any problem that wasn't advised by this channel, so we have this parallel thing, then I get together, we get better image, and also when you do traditional iterative reconstructing, so easy iterations, you do exactly the same thing, but when we have this unripe, say, ten part, each part you do training, then the easy iterations, you have different parameters, so iterative reconstructing is no longer a very, very strict fixed step, so through training easy studies can do something different, so some good features and good knowledge could be reflected by learned parameters, so this do things like this, and we achieve good performance in a typical application called SBAX Data CT, in medical CT, nowadays you scan data acquisition system around the patient, so each circular scan you collect roughly 1,000 projection views, let me say, what is the use of data using 1,000 projection views, it's just used 64 views, we have good reconstruction, so this is the reference image you reconstruct, it is all the projection, so 100,000, and if you only use 64 views, so filtered back of the projection, wouldn't look good, okay, and iterative reconstruction algorithm, like see, if you look carefully, blocky artifacts, not perfect, but based this network we call it Learn, learn the expert assignment based reconstruction network, so we make this special brand name Learn, it got better results, you see this learn result, you see the this darker region pretty much consistent to the reference, you see this traditional iterative algorithm, we have blocky artifacts as shown here, so overall we did the systematic evaluation in terms of quantitative measurement, and readers studied by radiologists, and we show learning masters, and this learn method work out performed competing masters, why is it an analytic master such as the filter back projection, or popular iterative master, or comprised sensing master, so this is example about reconstruction, for example, let me give you another example, I call it x-tereo tomography, or x-tereo problem, and then in your CT like this, you learn, that's the meta-artifacts, so you have meta-artifacts somewhere here, iterative block x-3, so the middle portion of the data is amazing, so what you measure are only on two parts, the green data part, you know, so we want to do image reconstruction, this has been a long standing problem, still no perfect solution, so masters do the collaborators wrote this review article talking about meta-artifacts reduction, we published in last year about 40 pages, a lot of people, but no perfect solution yet, and last year we used the machine learning in collaboration with Global Research Center, and in data domain and in image domain, so use machine learning to try to reduce meta-artifacts, so this is a new solution, if you have a meta here, and the meta block x-3, right, meta very dense, so actually cannot penetrate easily, so actually block the traces of your half-day cycle, no information available, then you cannot use filter-back projects and algorithm, and what about we failing the data in the guide, we can use linear interpolation, you can use cubic interpolation, and we know all these interpolations are not smart enough, wouldn't give you good results, not with machine learning, we train a neural network with big data, so we do smart interpolation, we're failing the data, then the meta-artifacts reduction problem becomes simpler, now you're failing the data, you just use filter-back projects and algorithm, you can do good reconstruction, okay, this is data domain, and the next I show you image domain master, okay, how you use image domain master to do meta-artifacts reduction, so this is also example, try to solve a problem, even you use machine learning master, you are not limited to a single master, the next master is image domain master, and we propose a super-ority principle, you just say, okay, you want to do meta-artifacts reduction, there are hundreds of papers already published on meta-artifacts reduction, now I'm perfect, but still you have very good method, okay, we select very good method, say, for meta-artifacts reduction, and we select good method, we do reconstruction, we guide the results better, but not perfect, okay, then we take the data of art result as input to a neural network, then we train the neural network with a gold standard, what you really want, you imagine you have a phantom here, this model, and you insert a metal different places, you do meta-artifacts reduction, not very perfect, then you know the gold standard, then you use gold standard and imperfect image, this higher data, together, train neural network, you have many, many tires, you have label the big data side, you train neural network, you get better result, better means you move little closer to the gold standard, and while it is starting point, you starting point is current master, data of art, so you wouldn't damage anything, you get a decent benefit with neural network, so this is the neural network, my student put together, so see multiple layers, you have rack linear unit, you use means, I draw lots, so very typical thing I mentioned, just many parameters, ten layers, you do the training, so we got good results, and the paper will be presented, I think July in Utah, so use machine learning to do meta-artifacts reduction, and even in image domain, the master is not unique, and this paper from Professor Hong Yong Yu's live, and their work flow is something like this, so they input into neural network, it's not only one image, you have different master to do meta-artifacts reduction, some even very simple master, like this, you use linear interpolation to fill in the gap, we know it's not perfect, but he still take data as one of many input, so you have original uncorrected image, you have some corrected, some simple way corrected, multiple images into the neural network trained by the gold standard, then put the results together, so this complicated work flow, so you see this original image, it was linear interpolation, I told you linear interpolation, it's not very good, you see this side of the streaking artifacts, you use some data of art result, you still have this dark strep, so many things to put together, this final input look is very, very good already, so this is just a clinical example, and the last example, give you it just opposite to the exterior tomography, and we divide the interior tomography, so just the opposite interior tomography, so you do not have data on both sides, you only have data in the middle, so you have a reason of interest to master data, associated with x-rays, going through this reason of interest, so we have multiple ways to do iterative reconstructions, to reconstruct small reason of interest, so in this case, we only send x-rays through this white circle defined the reason of interest, and we can get a very good reconstruction result, then the public think about eight years ago, so we got some results, and in most challenging cases, when the ROI is very small, the result will not be that good, so based data driven machine learning method, and we can correct for off-site, so this is another example, so a lot of details, we do not have time to cover all of them, so this is so your different ways to use machine learning for tomography reconstruction in several settings, so now let me give you second summary, and the deeper reconstruction, use deeper neural network for tomography reconstruction, so previously you have normal data side noise, it is not high, and you can use for example, filter the back projection algorithm, you perform an analytical reconstruction, you get good image, that's all the good days, and

now x-ray CT becomes so successful, many many scans, so people start worrying how about the dose, there are some dose, and this is about dose limitation, and low dose CT, so on, so to address the public condition and potential risk, if you have too much release in dose, you are subject to cancer risk, genetic damage, potentially, the chance, and I argue that it is not small, the natural background simulation, nowadays is comparable to a CT scan, under the benefit of far overweight the problem, but we still want to use small amount of release as feasible, so that's as less as possible, so we do low dose CT scan, then we need to use iterative reconstruction, and we got results somehow comparable to normal high dose reconstruction, so this is current status, so present situation, so past, and now, and what's the future, future we say, you can use machine learning, so you have low dose data, and in data domain you can use machine learning to input data, so you still have data size, and from this data size you could still utilize the algorithm, you already developed, analytic or iterative, so once you have image domain, you can use deep neural network smart processing, you get better sync, like denoising, I showed you, and all, and optimally, you go directly from raw data to images, so this is what I mean by deep reconstruction, so this is the future topic, and we are actively working on this, so this is the second brain home point, remember what's the first point, coding versus learning, the second point is deeper reconstruction, so feel there is a moving from analytic iterative reconstruction to data driven, learning based reconstruction, so this is trend, and say my group no longer actively research on iterative reconstruction, we do machine learning, dedicated to machine learning, but that does not mean we just the rule all these away, and as I explained in the learn network, we really combine the water we learn the iterative reconstruction into the network, as the imaging researcher, we have the knowledge and the insight to build most relevant network, if we just this card all our experience and knowledge, and other people like computer scientists would be able to do same thing, right, so you don't want to go that far, so this is brain home point for the second part, and these same slides, and while the site in planetary space, I think so called RISA meeting, and radiological society, no America, and this meeting is very huge, I think about 50,000 people, huge meeting, so you see the whole radiological society is moving towards machine learning direction, so last part, we have about 15 minutes, we finished last part about future imaging, and even more than what I explained, so this is the deeper we come to the future, but let me give you a broader viewpoint about future of imaging, so I wrote a debate article with my collaborator about machine learning, we will transfer radiology significantly, within five years, this is the first debate, we published about a year ago, and this article selected as a pop-10 by the journal, so many download people interested, the second debate published idea in the second debate article, entitled radiomics in lung cancer, time is here, let me explain this idea, what is radiomics, radiomics really, let me explain, so you have medical images already produced, this starting point is indeed images, you have images that really are made, this word really are made means you just analyze the image, you generate many, many features, like text feature, histogram and many features, like Fourier transformation, different pass filter, results, you just generate many features, the number of features is so huge, so the radiologists cannot possibly work with these features, so you need a computer program to do feature analysis, then based on this feature you let the computer program decide if you have a tumor or not, so this is from image to feature, machine learning can help to do that, now it's very hot topic several top people, the public is showing such machine learning based radiomics or deep radiomics, can beat human performance, another point of view I suggest, is something not just radiomics, I put letter W, I call it radiomics, so the argument goes like this, so these images, then you do features, so this domain is not seen by the water, just explained deep radiomics, this part, so I say one thing, and we need
to pay attention, I think we shouldn't just be satisfied with this part, the analysis should go back to the data domain, so if the data is perfect, then the data give you all four data information, so from data to image, then you reconstruct image, this will be ideally a invertible transform, so that means you can go from data to images, and from images you can go back to data, everything is perfect, you're multiple, so in that case you do image analysis, based on image or you do all the analysis, based on data, there will be no

difference, because the data and the image are equivalent, okay, I mentioned nowadays, and we use the load those CT, the data becomes very noisy, so you use iterative algorithm, you apply certain regularization means, so after that once you have images reconstructed, can you go back to your data domain, no, so this process is no longer invertible, there are no longer equivalent, in other words, the information, some information is lost in this process from data to images, therefore we think we can do image reconstruction, using representative algorithms, in multiple ways, so from data you reconstruct image one, use algorithm one, you do reconstruction tool, using algorithm tool, you do multiple reconstruction, this is consistent to the observation, many groups keep publishing paper on image reconstruction, or even in image processing, keep publishing paper on image segmentation, or as detection, and the easy paper saying my algorithm is better according to this data side than others measure, so what's the deal, everyone claim they are the best, I think they are not liars, they can select the data side for certain things, their algorithm indeed performs better than others, but the problem is that all these algorithms are neither of the algorithms solve all the problems, they are complementary, so if we let them reconstruct images, put all the images together, so we will not lose any significant information, so that's the idea we say from data side, you really need to use complementary algorithms, these algorithms like a unique way to extract some unique features, you use multiple algorithms, all unique features to put together, you are not subject to loss of information, so this is the precise from data to multiple copies of the complementary images to final features, we will give you more information, the whole thing can be streamlined because you go from raw data, you still have other information, so I call it raw data mix, so this is end to end work flow, so you can optimize the whole work flow and promising the best performance, okay, the last five minutes, let me give you my latest idea, I presented several, basically go, I took a school of engineering, school of engineering asked faculty members to brainstorm ideas which could potentially form a research center, so I think about ideas, this is the best idea I could come up with, so think about auto driving vehicle based foldable tomography analytics, like raw data mix, robots, so I call it a, so this idea can just a few minutes briefly explain, a few years ago we designed a very cheap low cost city scanner, all the off magnitude cheaper, and in my class this is one of the homework, so we can make cheaper, so this is one thing, so we have some track record, we are collaborating with global research center, and we try to develop a slame city scanner, this scanner can be put in supermarkets, CVIS, and you walk in there, and the load those easy operations, you do choose the screening, my biography, bone quality evaluation, and many things, so this will be very easy, right, so this is another component, third one I mentioned auto driving car becoming available, maybe next few years, yet another component, and robotic, both dynamic, dynamic, it's a good company, and they develop robotic, like four legs walking around, very stable, so all these things put together, so the reason is to have a future image in auto driving car based small machine, not limited to x3, and if you have need like you call Uber driver, the car comes to you, then you just have some robotic system, doing imaging analysis, so everything, this will be I think the future of imaging, so this represents convergence of multiple high-tech, and we are familiar with tomography imaging sensing, but also involves an electric, and the robot control driving, and telecommunication, internet, so all these things, is water required by NIsF, we will figure out the topic, you need to represent multiple disciplines, so this is the argument, also the technique should be classified to be valid data, the dimension data, the multiple types of bytes, why we bother to build robotic car, like I described, and water would be application, I say first, medical imaging will be easier, it just make a phone call, not only easier, it will be, could be a major way to cut a cost, the medical imaging not very accident, because you need a human, the scanner is in shielding room, you need space, but if you have a car, that's like Uber is cheaper than regular taxi, so this is one argument, and also we have occasionally natural disaster, so patients are here, you don't want to move, you really want to fly, or drive your imaging equipment to the spot, the battlefield, the terrorist attack, so you need that, also infrastructure imaging, you have building bridge, you need a more imaging capability, around features, structures, of interest, and space exploration, like Mars, you want to sign some imaging device, not like hospital clinic, you sign the imaging device, this is robotic, so this is future technology, I think that is correct, I do not know how soon it will come, so this idea seems very natural, to me, components, ideas, relevant ideas are flowing around, so the last one, summarizing class, one more minute, I think internet of service, no right now is internet of things, and my view next is internet of service, so civil legislation, spend by both mobility and connectivity, multiple examples, usually you go bus station, train station, and the private cars become very popular, Uber tax, so the entertainment, you go to cinema theaters, you have home theaters, now I watch video on phone, so very easy, the same thing happens, the imaging scanner, tomography scanner, nowadays center, in hospital clinic, in radiology department, and sometimes you have imaging centers, multiple scanners in it, you doctor, just a show or work order, you go there, but the future, the imaging service will come to you, come to seeing, answer about, so that's something I think, very exciting, so much for today, okay, so I could say, I'm in a position, this is on SkyGew, similar to first and second, and I will email my class about a few guidelines, so that's all for today, thank you, okay.