<Undergraduate Thesis: EEE-4610-01-00>

Deep-Learning Based Surface Profiling with Defocused Interferometric Plasmonic Microscopy Imaging

Jongha Lee
School of Electrical and Electronic Engineering
College of Engineering
Yonsei University

<Undergraduate Thesis: EEE-4610-01-00>

Deep-Learning Based Surface Profiling with Defocused Interferometric Plasmonic Microscopy Imaging

Thesis Advisor: Prof. Donghyun Kim

A thesis submitted in a partial fulfillment
for the senior Undergraduate Creative Independent Study's requirements

June 2022

Jongha Lee
School of Electrical and Electronic Engineering
College of Engineering
Yonsei University

감사의 글

먼저 김동현 교수님께 감사의 말씀을 드립니다. 이 졸업 연구와 더불어 연구실 학부생 인턴 기회를 허락하여 주셔서 제가 수준 높은 연구에 참여할 기회를 주셨습니다. 교수님께서 소중한 시간을 내셔서 연구 지도와 피드백을 주신 덕분에 더 좋은 연구를 할 수 있었습니다.

또한, 발표 심사와 평가를 담당해주신 김동현 교수님께도 감사의 말씀을 드립니다. 저는 이 졸업 연구를 수행하고 평가받는 과정에서 저희 학과 두 분의 김동현 교수님께 많은 도움을 받았습니다. 평가를 맡아 주신 김동현 교수님의 통찰력 있는 리뷰와 좋은 논평들을 통해 연구를 더욱 발전시킬 수 있었습니다. 덕분에 더 완성도가 높은 연구를 진행할 수 있었다고 생각합니다.

광생체공학연구실에 있는 모든 연구원들께 감사를 드리고 싶습니다. 특히 문귀영 연구원께 감사의 말씀을 드립니다. 제가 처음으로 연구를 진행하는 상황을 잘 이해해 주시고, 주제를 잡는 과정, 연구를 진행하는 방법과 같은 전체적인 연구 과정을 찬찬히

가르쳐 주셔서 큰 도움이 되었습니다. 첫 연구에 너무나도 훌륭한 멘토를 만나게 되어 큰 행운이라고 생각합니다. 또 연구를 진행하는 과정 중, 이현웅 연구원과 임성민 연구원께서 피드백과 실험 진행 방향과 방법에 대해서 조언을 주셔서 감사의 말씀을 드립니다.

상하이 교통대학에 있는 Hui Yu 교수님과 Jingan Wang 연구원에게도 감사를 드립니다. 연구 주제를 사전 조사하면서 의문점이 생겨서 연락을 드렸을 때 너무나도 친절하게 답변을 주셨습니다. 덕분에 연구를 수월하게 진행할 수 있었으며 연구에서 협력의 중요성을 느낄 수 있었습니다.

마지막으로 연구 외적으로도 도움을 주신 부모님과 가족에 진심으로 감사와 사랑을 전합니다. 항상 도움을 주고 필요한 받침이 되어서 늘 감사합니다.

Contents

Figure index

Table index

ABSTRACT

Deep-Learning Based Surface Profiling with Defocused Interferometric Plasmonic Microscopy Imaging

Surface Plasmon Resonance Microscopy is an extremely versatile imaging technique, with unique advantages such as selective imaging of phenomena near the film surface and increased sensitivity based on interferometric patterns. The interferometric patterns created by nanoparticles located on a film contains a variety of information concerning the film and the nanoparticles. However, traditional methods of analysis using the point spread function are unsuitable for the analysis of this information because of their size being far greater than the diffraction limit and their parabolic shape. Deep learning based analysis techniques are introduced to this work to analyze the images created by surface plasmon resonance microscopy. The results show that a conventional convolutional neural network can be used reliably to extract this information and surface profiling of the sample is carried out to prove this concept.

Key words: Surface Plasmon, Plasmonic Microscopy, Deep-Learning, Defocus Imaging, Surface Profiling

1. Introduction

Surface plasmons have been utilized for microscopy for decades because of their ability to create super-resolution images with relatively few restrictions on the requirements of the sample preparation. The advancements of nanofabrication methods and optical analysis techniques have further increased the possibilities of imaging techniques based on surface plasmons. One of the biggest advantages brought by the usage of such techniques is the ability to acquire label free images with high spatial and temporal output with a wide field of view. However, the spatial resolution is mostly limited to the 2-dimensions parallel to the sample surface [1].

Deep learning has gained interest in the past decade for its ability to analyze images. The success of the AlexNet network in the 2012 ImageNet competition [2] has shown the promise of deep convolution neural networks in the area of computer vision and research on their applications has advanced rapidly since its first proof of concept. The deep neural network architecture is naturally suited for recognizing patterns in the input image, making them a highly attractive method for extracting information from given images.

The parabolic shape of interferometric patterns caused by surface plasmon resonance microscopy and their large shape far greater than the diffraction limit prove to be a challenge for analysis using traditional methods. Deep learning can be used to analyze the received images rather than using traditional analytical methods. Combining the advantages of deep learning with surface plasmon microscopy, this work aims to measure the normal location of nanostructures by analyzing surface plasmon resonance images.

This work profiles the surface of gold films as a proof of concept for this approach. Gold nanoparticles and E-beam resist nano-posts are deposited on a gold film. The interferometric plasmonic images are taken of these nanostructures. The images are analyzed using a modified convolutional neural network for height regression. The resulting surface profiles are compared to the ground truth surface profile measured through atomic force microscopy.

2. Background

2.1 Surface Plasmons

Surface plasmons are a phenomenon that are known to occur at the interface between metal and dielectric materials [5]. This occurrence may happen at any interface where the real part of the complex refractive index changes signs. Surface plasmons are characterized by a coherent oscillation of the electrons at the interface with an electromagnetic wave. Metals naturally contain free electrons with the ability to oscillate, allowing this coherence to appear. The oscillation of electrons at the interface are called surface plasmons.

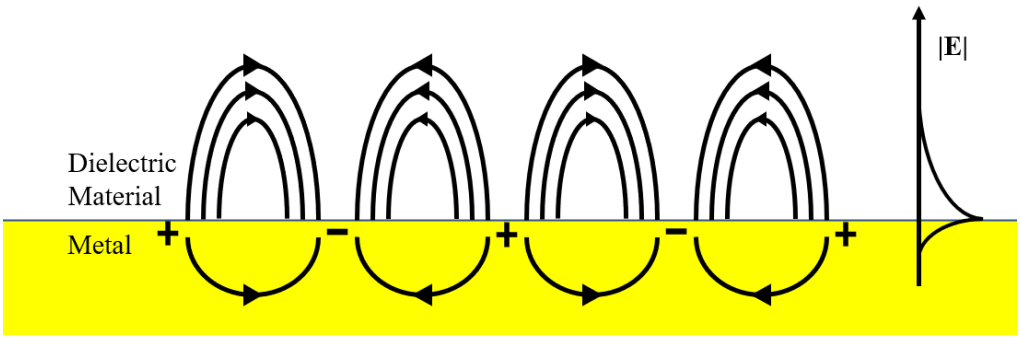

Figure 2.1.1. Creation of surface plasmons using incident electromagnetic waves. The waves exponentially decrease in the normal direction. The top material is a dielectric material, and the bottom material is a metal. The oscillation of free electrons in the metal creates this phenomenon.

Most surface plasmons are researched under the condition of the electromagnetic wave being incident on a dielectric material and metal interface while propagating from a prism. It will be assumed that surface plasmons is this work are created at the dielectric material and metal interface.

Surface plasmons offer interesting optical properties. The oscillations are caused by electrons, causing the wavelength to be significantly shorter than the wavelengths of electromagnetic waves with identical frequencies. A simple example of this can be achieved

using an HeNe laser with an output electromagnetic radiation with a wavelength of 633nm incident on a Si/Ag interface. The wavelength of the surface plasmon can be as small as 70nm, less than a ninth of the original electromagnetic wavelength [3]. The wavelength of the surface plasmon depends on many factors, such as the complex refractive index of materials and the thickness of the metal film.

Surface plasmons decay exponentially in the normal direction of the interface. This causes surface plasmon waves to become more sensitive to surface aberrations and ignore information about features that are far from the interface. Small changes in the surface of the metal can cause large changes in surface plasmons. Detecting these enhanced changes enables optical resolution better than the traditional diffraction limit of half of incident light wavelength.

2.1.1. Surface Plasmon Polariton

Polaritons are quasiparticles that are created from a strong coupling of electromagnetic waves with electric or magnetic excitation [5].

Surface plasmons under specific conditions are observed to achieve strong coupling with the incident electromagnetic wave. The resulting electron oscillation from this phenomenon is referred to as surface plasmon polaritons. Surface plasmons are confined in the normal direction. However, specific plasmon modes are able to propagate in the lateral direction when coupling is achieved between electromagnetic waves and free electrons within the metal. These surface plasmons are known to be propagating surface plasmon polaritons. The conditions required for this strong interaction depend on the wavelength and incident angle of the incident electromagnetic radiation. In most cases, the incident angle of the electromagnetic radiation is large.

2.1.2. Surface Plasmon Resonance

Surface plasmon resonance is a synonym for surface plasmon polaritons. Resonant oscillation of electrons is required in surface plasmon polaritons. Surface plasmon resonance stresses the resonant nature between free electrons and the electromagnetic waves in this phenomenon.

2.1.3. Surface Plasmon Polariton Equations

Surface plasmons polariton behavior can be derived from Maxwell's equations at the boundary of the dielectric-metal interface. The incident light and the surface plasmon polariton is modeled in Fig. 2.2.

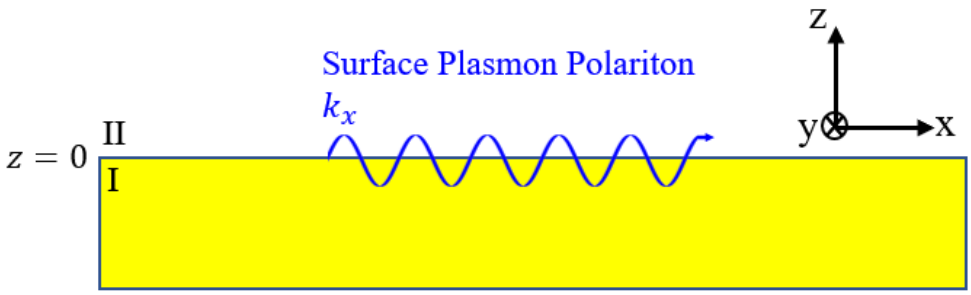

Figure 2.1.2. The surface plasmon polariton propagates along the interface. Material I is a metal and material II is a dielectric material. The surface plasmon polariton propagates on the interface with a wave vector $\boldsymbol{k_x}$.

The electric and magnetic plane waves incident on the interface in the transverse magnetic (TM) polarization can be expressed as follows.

$$E_{x,n}(x,y,z,t) = E_0 e^{ik_x x + ik_{z,n}|z| - i\omega t}$$

$$E_{z,n}(x,y,z,t) = \pm E_0 \frac{k_x}{k_{z,n}} e^{ik_x x + ik_{z,n}|z| - i\omega t}$$

$$H_{y,n}(x,y,z,t) = H_0 e^{ik_x x + ik_{z,n}|z| - i\omega t}$$

Where E and H indicate the electric and magnetic fields respectively. Cartesian coordinates shown in Fig. 2.1.2. are used. n is used to indicate the refractive index of different materials with 1 indicating the metal film and 2 indicating the dielectric material. k, ω, and t indicate the wave vector, angular frequency, and time respectively.

The electric field and magnetic field intensity relation can be found by considering plane wave propagation characteristics. The relation between the two magnitudes is shown below.

$$\frac{H_0}{E_0} = -\frac{\varepsilon_1 \omega}{k_{z,1} c} = \frac{\varepsilon_2 \omega}{k_{z,2} c}$$

ε is the permittivity of the materials, and c is the speed of light in a vacuum.

Using the fact that the electric flux density is 0 at the interface, the following relation concerning the normal wave vector can be derived.

$$\frac{\mathrm{k_{z1}}}{\varepsilon_1} + \frac{\mathrm{k_{z2}}}{\varepsilon_2} = 0$$

The relation between the normal and tangential wave vectors can be found by comparing the square sums inside each material. The relation can be expressed as shown below.

$$k_x^2 + k_{zn}^2 = \varepsilon_n \left(\frac{\omega}{c}\right)^2 \qquad n = 1,2$$

Using boundary conditions of the wave vector over the interface, the tangential wave vector of the surface plasmon can be found. The complex wave vector of the surface plasmon polariton is shown below.

$$k_x = \frac{\omega}{c}\left(\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}\right)^{1/2}$$

Where ω is the angular frequency of the surface plasmon.

If we assume that the angular frequency ω is a real value, permittivity of the dielectric material ε is real, the complex wave vector can be separated into the real and imaginary parts. The complex wave vector can be expressed as expressed below. The prime and the double prime each indicate the real and imaginary parts of a complex number.

$$k_x = k_x' + i k_x'' = \left[\frac{\omega}{c}\left(\frac{\varepsilon_1' \varepsilon_2}{\varepsilon_1' + \varepsilon_2}\right)^{1/2}\right] + i\left[\frac{\omega}{c}\left(\frac{\varepsilon_1' \varepsilon_2}{\varepsilon_1' + \varepsilon_2}\right)^{3/2} \frac{\varepsilon_1''}{2(\varepsilon_1')^2}\right]$$

Solving the equation for the transverse electric (TE) polarization gives us a surface plasmon of amplitude 0. The surface plasmon polariton exists only in the TM polarization and can only be excited by light in the TM polarization. This requires the controlling of the polarization mode of the incident light when the creation of surface plasmons is desired.

2.1.4. Localized Surface Plasmon Resonance

Surface plasmons refer to the phenomenon of free electrons on the surface of the interface oscillating at a particular wavelength. Localized surface plasmon resonance occurs when aberrations are existent on the surface of the metal. Nanoparticles on the surface of the metal provide free electrons that are easily polarizable. However, contrary to other electrons within the usual metal film, these electrons are confined to the particle volume. A coherent shift in the location of electrons caused by an electromagnetic wave causes a Coulombic restorative force from the now positively charged nanoparticle. An oscillation is created because of this restorative force. The oscillation of these electrons is named localized surface plasmons because of their confined nature.

The localized nature of localized surface plasmons polaritons gives them enhanced properties near the nanoparticles. The electric field near the nanoparticle is enhanced compared to other regions without nanoparticles, allowing for extremely sensitive detection near the nanoparticle.

The excitation of localized surface plasmon resonance can be finely tuned by changing the nano-particle shape, size, composition, arrangement, and incident electromagnetic wavelength and incident angle.

2.1.5. Gold in Surface Plasmons

Gold is one of most frequently used metals in work concerning surface plasmons [7]. Gold boasts a high level of chemical and physical stability, opening possibilities in biosciences. Furthermore, gold nanofilms and nanoparticles are observed to achieve resonance at visible – near infra-red frequencies, which makes them much more attractive for usage in optical experiments. This is caused by their permittivity coefficients and other noble metals such as silver are often used for surface plasmons as well.

The fabrication of gold films is also simple to implement. The usage of relatively simple fabrication techniques such as evaporation can be used to create a uniform film on a given surface.

2.1.6. Electron-beam (E-beam) Resist Structures

E-beam structures are often used in plasmonics for their versatility and flexibility in fabrication [10]. E-beam structures are created by using E-beam resistors and photo lithography to “carve” the resistors into the desired shape. The sample is coated with an E-beam resist material. The material is exposed to E-beams and depending on the material, the material exposed or not exposed to E-beams are removed during a process called developing. The remaining E-beam resist creates the desired structure on the sample [11].

E-beam structures are easy to manipulate because of their lithography process. The size, shape, orientation, etc. can all be controlled through the lithography process. The structure geometry sensitivity of surface plasmon resonance is able to take advantage of this versatility to create the desired resonance characteristics.

2.1.7. Plasmonic Microscopy

Plasmonic microscopy is a method of combining the enhancing ability of surface plasmons with microscopy to achieve greater resolution compared to when using a normal light as an illumination source. First proposed in 1988, plasmonic microscopy has proven to be a powerful label-free tool for sample analysis [9]. The enhanced fields are much more sensitive to changes at the surface of the film, allowing for high contrast imaging. Miniscule changes in location and refractive indices can be detected. Furthermore, the near-field enhancement feature of plasmonic microscopy means that phenomena that happen far away from the surface is automatically ignored [8]. Plasmonic microscopy can be used in a label-free environment, enabling the observation of a variety of samples.

2.2. Metal-Dielectric System and the Kretschmann Configuration

There are largely two different metal-dielectric configurations used to excite surface plasmons, the Otto configuration and the Kretschmann configuration. The Otto configuration is difficult to fabricate and analyze compared to the Kretschmann configuration. The current dominant method of creating surface plasmons is the Kretschmann configuration.

The Kretschmann configuration using gold and air as the metal and the dielectric material is shown in Fig. 2.3. The Kretschmann configuration has the prism and the gold film directly in contact with each other. The surface plasmons are excited at the boundary between the gold film and air. The prism-metal interface creates an evanescent wave, and the gold film acts as a spacing layer to excite surface plasmons at the gold-air interface. The surface plasmon polariton propagates along the surface of the gold film. The incident angle of the electromagnetic wave is large to cause surface resonance.

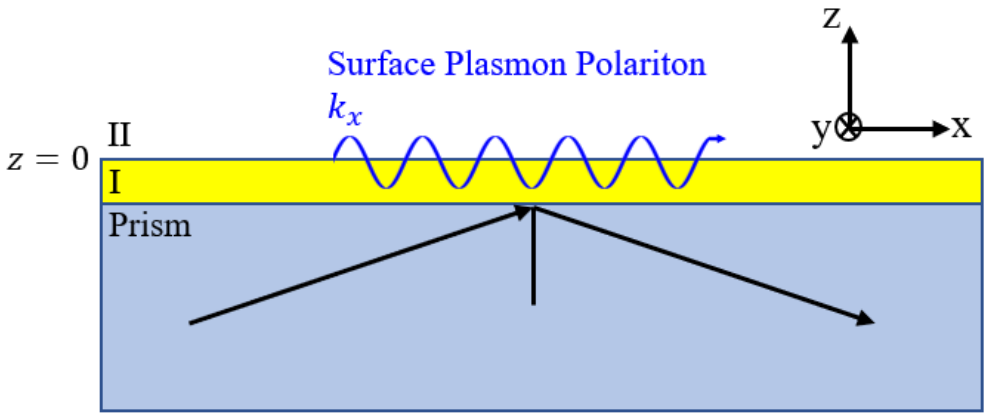

Figure 2.2.1. Creation of the surface plasmon polaritons using the Kretschmann configuration. The incident electromagnetic wave on the gold film creates a surface plasmon polariton on the gold and dielectric interface.

This configuration presents a significant decrease in fabrication difficulty while preserving the ability to excite surface plasmons. The evanescent wave travels through the air (dielectric) layer and excites the electrons in the gold film. This evanescent wave creates surface plasmons at the air-metal interface towards the direction of the prism. However, such a configuration creates difficulties in implementation as the gold film needs to be within a nano-scale distance from the prism for the evanescent wave to not decay a significant amount.

2.3. Interferometric Plasmonic Microscopy (iPM)

Interferometric plasmonic microscopy is a method proposed to further improve surface

plasmon resonance microscopy with higher sensitivity [12]. This type of microscopy is carried out using the Kretschmann configuration shown in Fig. 2.3.1.

The incident electromagnetic wave illuminates the gold film at a large incident angle. Reflection occurs at the prism-gold interface. Nanoparticles on the gold film surface scatter the incident light. The scattered light is the leakage radiation of localized surface plasmons. When the incident angle of the electromagnetic wave is near the surface plasmon resonance angle, the leakage radiation creates a propagating objective wave. The objective wave and the reflected wave are coherent and combine to form a hologram. This image can be observed by a camera, and the interferometric pattern can be analyzed to obtain information about the scattering object.

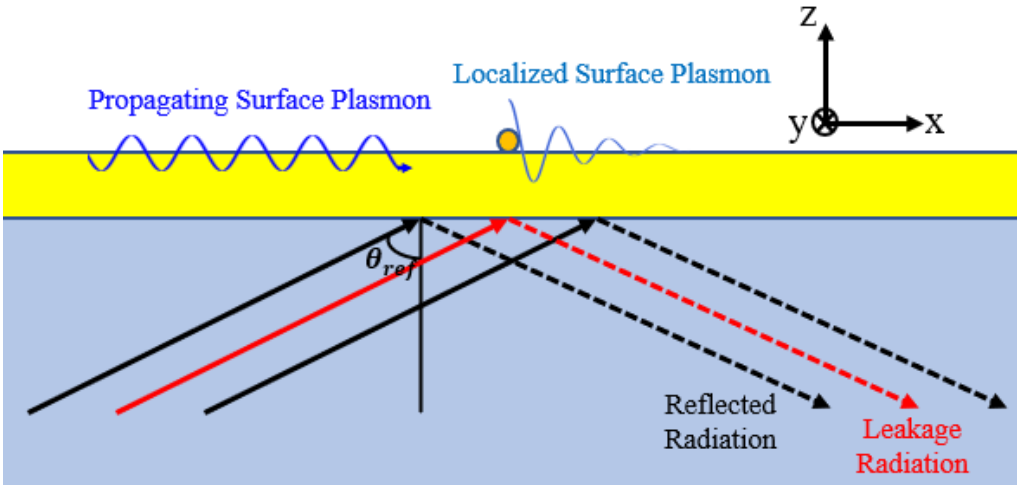

Figure 2.3.1. Electric fields at the gold film. The propagating surface plasmons and the localized surface plasmons at the gold interface interact, creating the leakage radiation. The reflected radiation from the gold film and the leakage radiation combine to create the final output electromagnetic wave.

This object scattered light and the reflective light from the gold film is observed using an electron-multiplying charge coupled device (EMCCD) camera. A background reflective image is separately recorded and is subtracted from the image created by a gold film with objects, creating a differential image. The differential image shows the small intensity fluctuations caused by scattering from aberrations located on the surface of the film. This information can

be used to locate the position of the particle with extremely high spatial resolution.

2.3.1. Observed Intensity of iPM

The basic principle of iPM is common-path interferometry where light from different sources share a common path, creating interferometric patterns. There are two sources of electromagnetic waves for the received light. The reflected light and the leakage radiation of surface plasmons. The leakage radiation caused by the nanoparticles creates an objective wave in a small area around the surface plasmon resonance angles. The objective wave and the reflective wave interfere, creating the observed electromagnetic wave distribution.

The detection device at the other end of the configuration detects the light intensity at the film surface. The leakage radiation and the reflected wave combine to create the final observed intensity. The formula to calculate the detected intensity is shown below.

$$I = |E_R + E_O|^2 = |E_R|^2 + |E_O|^2 + E_R E_O^* + E_R^* E_O$$

E_R is the reflected wave and E_O is the objective wave caused by leakage radiation. The magnitude of the reflected wave is much greater than the leakage radiation. Therefore, from the detected intensity graph, the absolute value of the objective wave squared can be ignored. The reflected wave is much larger than the other elements in the detected image; however, the reflected wave contains information about the reflectance of the gold film which is constant over the entire film and contains no information about the nanoparticle on the gold film. Therefore, during post-processing, the detected intensity graph is subtracted by the intensity of the reflected wave. The output of the post-processed detected image can be expressed as shown below.

$$I' = |E_R + E_O|^2 - |E_R|^2 \approx E_R E_O^* + E_R^* E_O$$

The electric fields at the interface can be simulated by adding the leakage radiation and the reflective radiation at the surface. The reflected wave can be analytically calculated using the wave equations for the incident electromagnetic wave. Setting the amplitude of the reflective wave as 1, the reflective wave can be expressed as follows.

$$E_R(x,y) = e^{-ik(y\cos\theta + x\sin\theta)}$$

Θ is the angle at which the incident electromagnetic wave propagates relative to the x axis on the x-y plane. x and y indicate the location of interest on the metal film. The simulated results of the reflected wave are shown in Table 2.3.1.

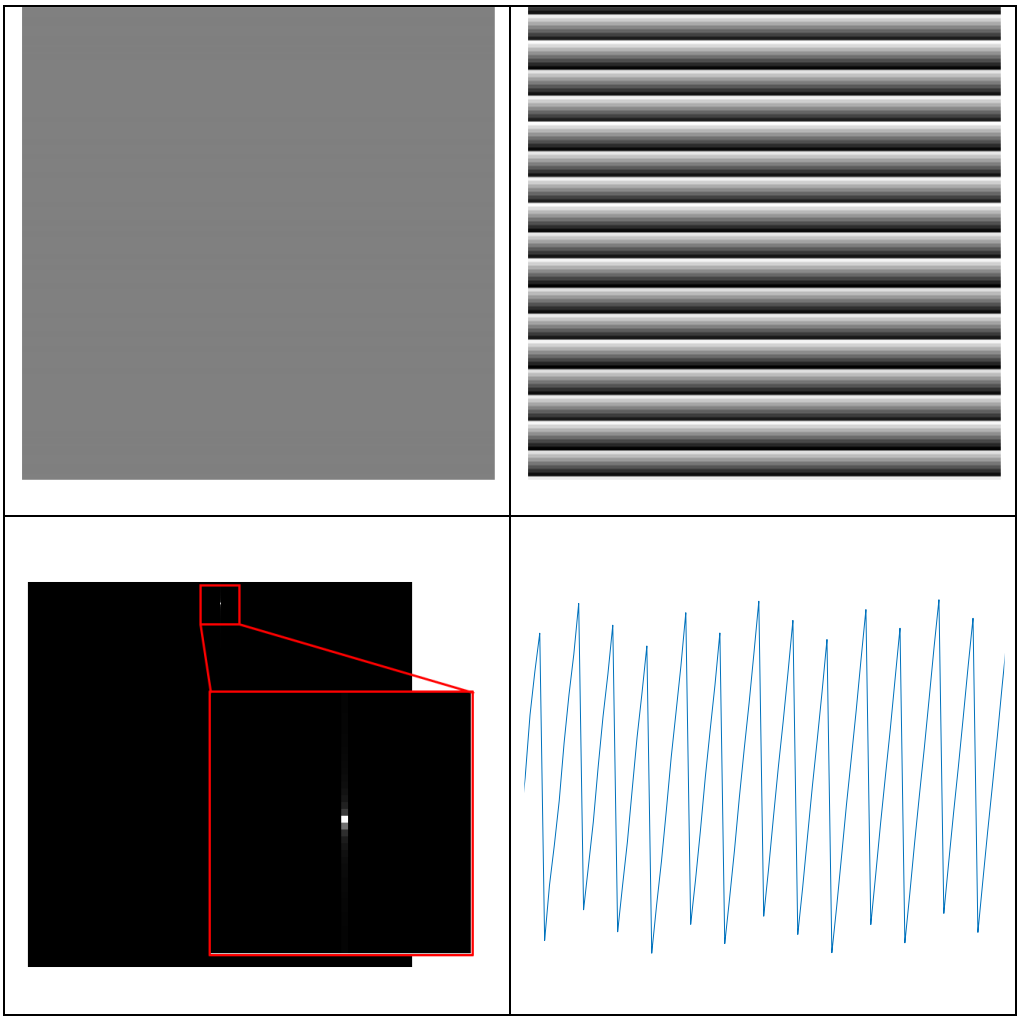

Table 2.3.1. Reflected wave from the gold film. Top left: magnitude of the reflective image. Top right: phase of the reflected image. Bottom left: The reflected wave in k-space. Bottom right: The phase of the reflected wave over a vertical slice.

The reflective wave has a constant magnitude at all parts of the film, with the phase of the electromagnetic wave increasing from the bottom of the image to the top of the image and oscillating.

The leakage radiation is the objective wave of the nanoparticles on the gold film surface. This wave can be modeled using the equations used to describe surface plasmons in previous sections. At the surface of the gold film, the surface plasmon resonance equations determine the electric field distribution. The objective wave of the electric field at the interface caused by a single nanoparticle can be expressed as shown below.

$$\mathrm{E_{Os}}\big((\mathrm{x},\mathrm{y})\big|(\mathrm{x_0},\mathrm{y_0})\big) = \alpha \mathrm{E_R}(\mathrm{x0},\mathrm{y_0}) \cdot \mathrm{e}^{-2\mathrm{k}''\sqrt{(\mathrm{x}-\mathrm{x_0})^2+(\mathrm{y}-\mathrm{y_0})^2}}\mathrm{e}^{-\mathrm{ik}'\sqrt{(\mathrm{x}-\mathrm{x_0})^2+(\mathrm{y}-\mathrm{y_0})^2}}$$

The objective wave magnitude is linearly related to the reflective electric field magnitude at the location of the nanoparticle. The objective wave is amplified as a stronger electromagnetic wave is incident on the nanoparticle which creates the objective wave of interest.

Cartesian coordinates are used to define the location on the surface of the interface. $(\mathrm{x_0},\mathrm{y_0})$ are the coordinates of the nanoparticle. We assume that all phenomena happen at the surface and ignore the normal component. The electric field objective wave at the surface of the interface is equivalent to an exponentially decreasing electric wave with respect to the distance from the nanoparticle with spatial oscillation. The objective wave created by nanoparticles are a cylindrical wave. The rate of oscillation and exponential decrease in magnitude is determined by k' and k'', which are determined by the complex permittivity of the materials used for the experiment.

The objective wave with the coordinates changed to their vectorized forms can be expressed as shown below.

$$E_{Os}(r|r') = \alpha E_R(r') \cdot e^{-2k''|r-r'|}e^{-ik'|r-r'|}$$

r' is the location of the nanoparticle and r is the location on the interface.

The resulting objective wave created by the scattering caused by a nanoparticle on the surface of the film is shown below.

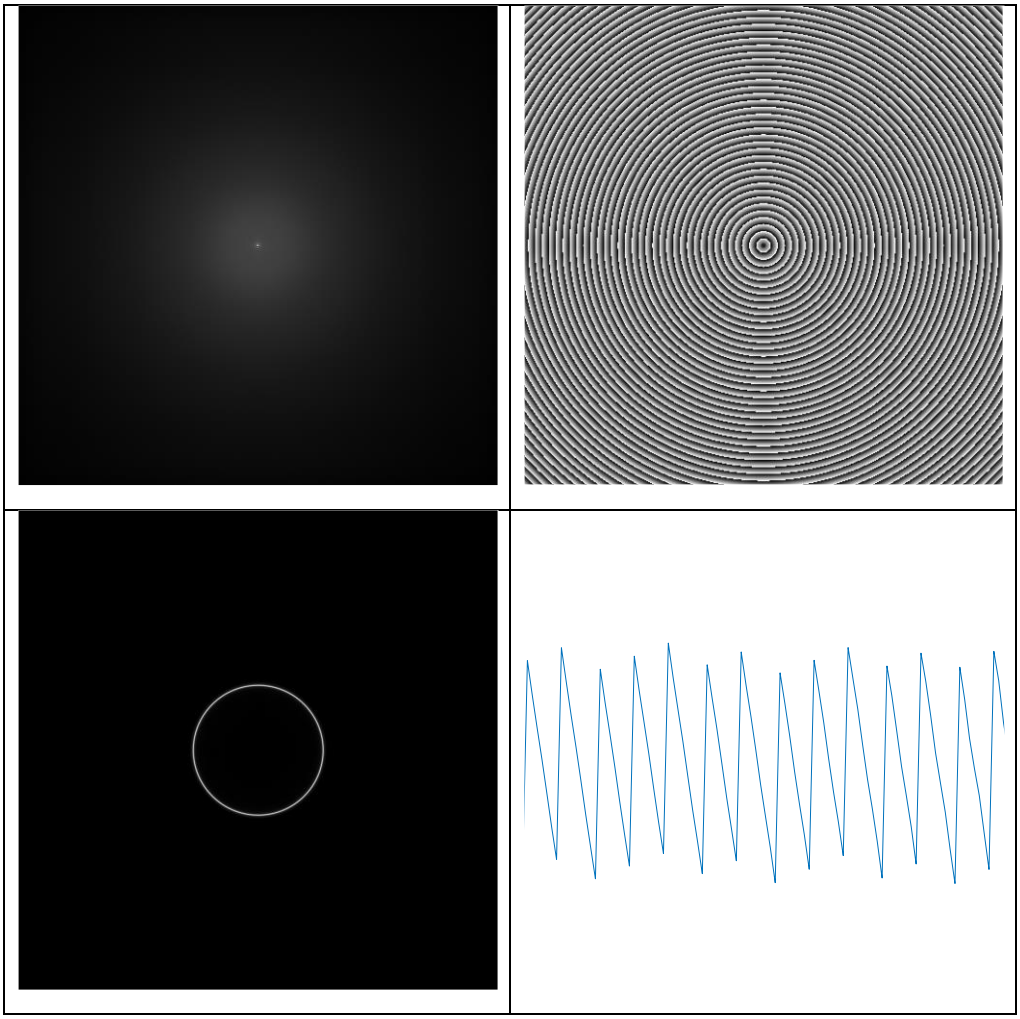

Table 1.3.2. The objective wave. Top left: magnitude of the objective wave. Top right: phase of the objective wave. Bottom left: objective wave in the k-space. Bottom right: phase of the reflective wave in a vertical slice.

It is assumed that the nanoparticle is located at the center of the image. The objective wave is seen to spreading from the center of the image. A cylindrical electromagnetic wave can be observed with oscillating phases and decreasing intensity as the wave propagates from the center.

The above derivation assumes that the nanoparticle is small enough to be thought of as a point scatterer. The coefficient α is the constant that represents the polarizability of the nanoparticle. A higher value of α indicates an objective wave with a higher magnitude from the

nanoparticle. The time component is ignored in this formula. The time component if added will be expressed as a complex phase shift depending on time and the frequency of the incident electromagnetic wave.

The combination of the reflected wave and the objective wave creates the final observed interferometric pattern. However, if there are multiple nanoparticles on the surface, the objective wave can be simply found by summing the individual objective waves.

$$E_O(x, y) = \sum E_{Os}$$

Using the formula above for the post-processed image, the final recovered interferometric image is in Table 2.3.3.

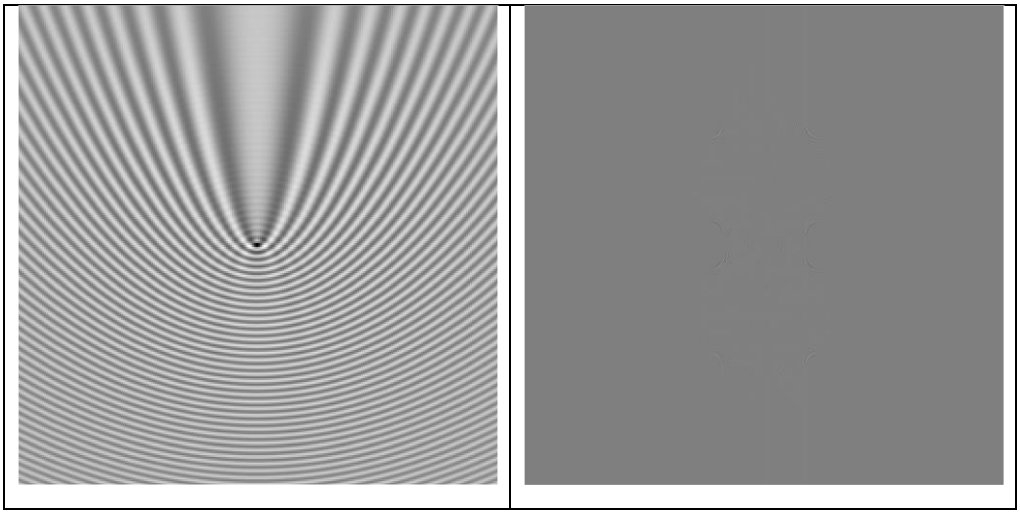

Table 2.3.2. Example of a simulated interferometric image. Left: received interferometric image after processing. Right: Post process image in the k-space.

The interferometric pattern creates a distinctive parabolic shape. In the k-space, two circles that are tangent to each other can be observed.

This interferometric pattern is created by a point scatter and is also created regardless of the scatter point material. However, depending on the complex refractive index of the scatter point material, the objective wave has a different phase. This leads to a change in the final interferometric pattern. This can be observed in Table 4.

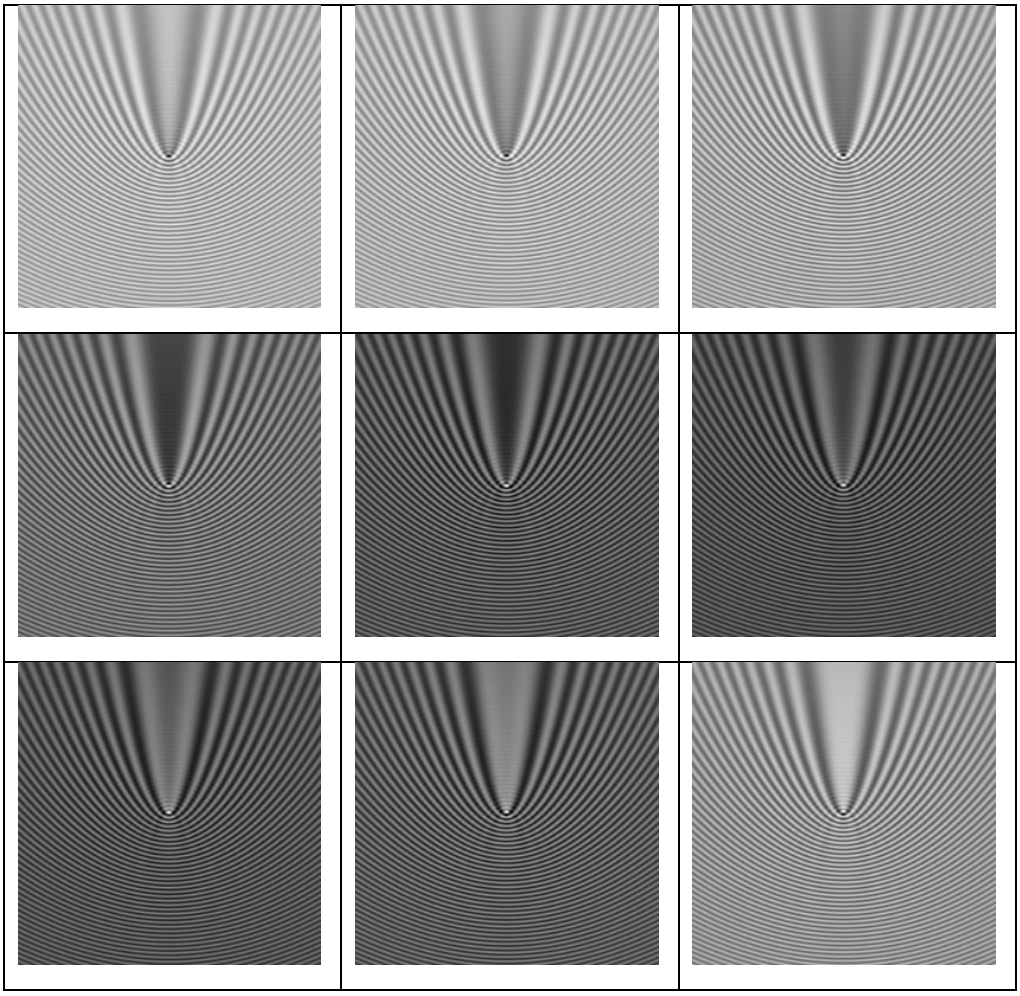

Table 2.3.3. The created interferometric pattern depending on the phase of the point scatter object. The phase of the point scatter object is dependent on the complex refractive index of the point scatterer. The phase of the nanoparticle is from top left to bottom right: $-\mathbf{0.8}\boldsymbol{\pi}$ to $\mathbf{0.8}\boldsymbol{\pi}$ with an increment of $\mathbf{0.2}\boldsymbol{\pi}$.

The created interferometric patterns all have the same shape. However, the gray scale is seen to be shifted. The phase shifting nature of the interferometric pattern depending on the phase of the point scatterer can be observed.

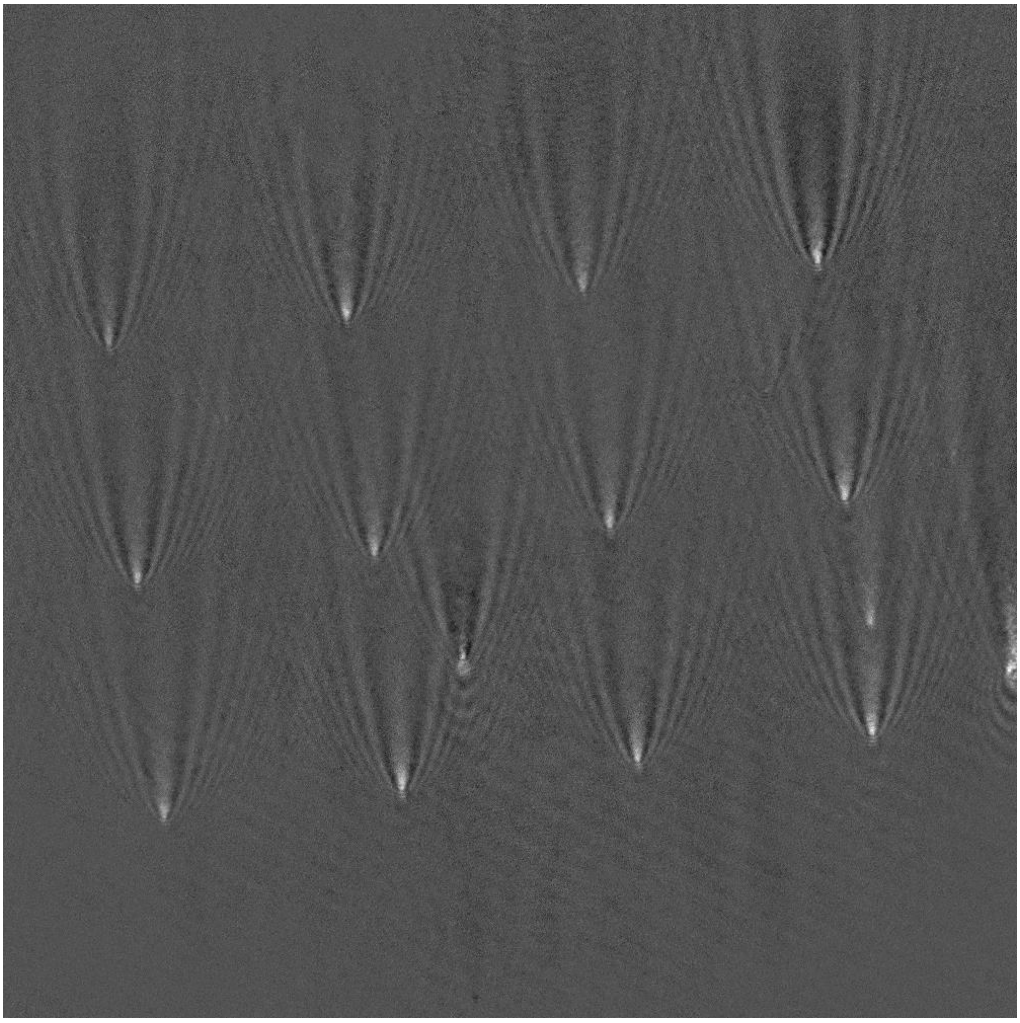

Figure 2.3.2. A physical image taken of interferometric plasmonic microscopy. The parabolic shape resembles the simulated images.

A physical image taken using interferometric plasmonic microscopy can be seen in Fig. 2.3.2. The distinct parabolic shape created by the images are extremely similar to the real-world image created by a physical implementation.

2.3.2. Intensity at the Imaging Plane

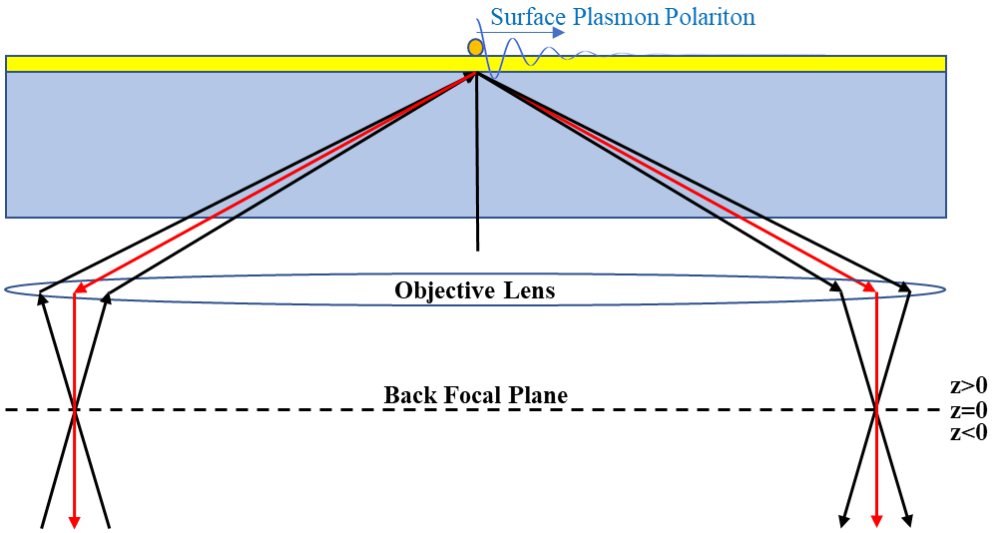

Figure 2.3.3. The movement of the back focal plane. When z=0, the back focal plane observes the electromagnetic field at the film surface. Movement of the back focal plane causes the collection of defocused images.

The equations from the previous section only apply when the back focal plane of the detecting device is focused on the gold film surface. The movement of the back focal plane with respect to the location of the nanostructure and gold film causes the observation of modified object and reflective waves [13]. The observed images when the focal point is moved are called defocused images because of the movement of the focal point away from the gold film surface.

The detected intensity at the imaging plane can be calculated by considering the location of the back focal plane relative to the imaging plane. The formula for the objective wave observed at the imaging plane depending on the varying position of the back focal plane can be modified as shown below.

$$E_O(r, r', z) = \alpha E_R(r') \cdot e^{-2k''|r-r'|} e^{-ik'|r-r'| + \sqrt{|nk_0|^2 - |k_{sp}|^2}\, z}$$

z is the movement of the back focal plane, detailed in Fig. 2.3.2.

This allows for the creations of different interferometric images depending on the z location of the nanoparticle relative to the back focal plane. The retrieval of information concerning the location of the gold film with respect to the focal plane is possible, providing information on the normal component of the objective wave. The different interferometric patterns depending on the z value is in Table 2.3.5.

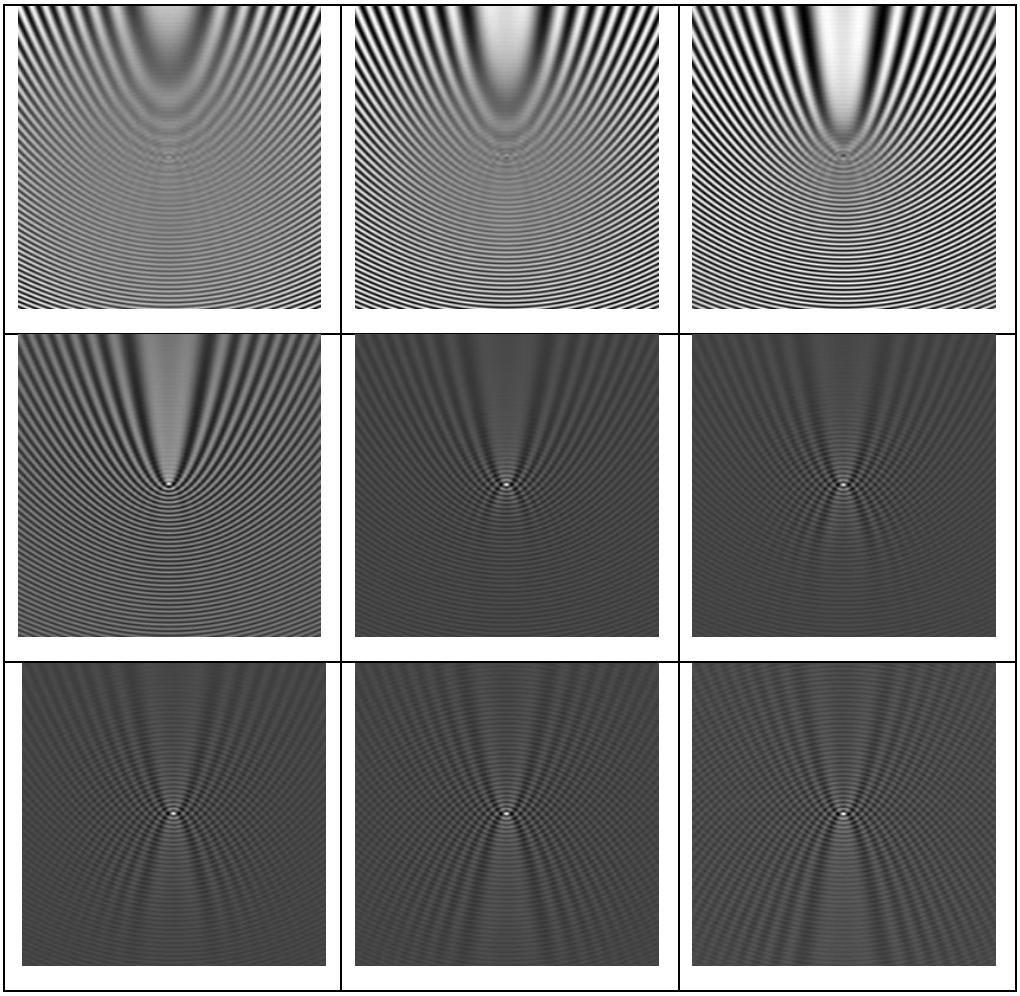

Table 2.3.4. Different interferometric patterns caused by movement of the back focal plane. From top left to bottom right, the value of z is -3 to 5μm with an increment of 1μm. As the back focal plane is moved toward the film, the interferometric pattern is changed from a parabolic shape to an X shape.

Multiple different methods are possible in changing the value of z in the formula above to change the behavior of the objective wave and create different interferometric patterns. The changing of the back focal plane by changing the focal point is possible but changing the normal location of the gold film and nanoparticle or changing the film thickness is also able to change the objective wave and the interferometric pattern.

Previous methods of determining the location of the nanoparticle relative to the back focal plane have included methods such as measuring the variance of the interferometric image compared to the overall average of the post-process picture [11, 14]. The results of this solution provided high resolution but required a molecular tether to the film surface. The label-free general-purpose advantage of surface plasmon microscopy is lost.

2.4. Numerical Aperture

Numerical apertures are the size of the holes in objective lenses that allows the passage of light such that they can be detected. However, numerical apertures are limited in size, and this creates an optical effect similar to a 2D low pass filter. The incident electromagnetic waves on a numerical aperture can be thought of as an aggregate sum of multiple electromagnetic waves with differing wave vectors. This is based on the principles of Fourier optics, where any wave can be thought of as a sum of multiple plane waves. The size of the numerical aperture determines the cutoff of the high frequency components incident on the numerical aperture. A larger numerical aperture is able to contain more information.

2.5. Atomic Force Microscopy (AFM)

Atomic force microscopy is a type of scanning probe microscopy [18]. A mechanical probe called a cantilever scans the surface of a given sample, measuring the response of the sample at an extremely high resolution. AFM does not use optics which allows for the creation of images with resolution far higher than the optical diffraction limit. The AFM imaging system can be used for topographic imaging of a sample. This topographical map created by the AFM is extremely high quality and can be used as the ground truth image when comparing topographical maps created by different methods.

2.6. Deep Learning

Deep learning is a branch of machine learning where neural networks with multiple hidden layers are used for learning tasks. Such methods are especially suited for data intensive tasks such as image analysis. With the advancements of analysis techniques and computational power brought by hardware and software development, deep learning has been proven to be an effective method of complex data analysis.

2.6.1. Convolutional Neural Network

After the success of AlexNet in the 2012 ImageNet Large Scale Visual Recognition Challenge, convolution neural networks (CNN) have gained interest as an effective method of image analysis. The introduction of large data sets as well as the increased computational abilities through hardware advancements have made CNN a widely used method in image analysis. Their natural ability to preserve spatial information makes them especially suited for classification and pattern recognition.

CNNs are multi-layer perceptrons with a structure similar to the connections of the human visual cortex. A single neuron is responsible for detecting changes within a small receptive field. The feature maps in a CNN respond to changes within a certain receptive field in the previous layer. The size of the receptive field accumulates for feature maps in deeper layers, allowing for deeper layers to respond to changes in a large area of the original input image. Shallow layers are able to contain low level features such as edges while deeper layers are able to contain high-level features such as image-wide patterns.

2.6.2. VGGNet

The VGGNet, first introduced in 2014 showed that deep neural networks are capable of extracting high-level features from images without explicit modeling [16]. This ability to extract high-level features was shown to improve with the depth of the model architecture, proving the need for deep networks in pattern recognition.

2.7. Noise

The existence of noise in real-world systems is unavoidable. The errors caused by sensors during acquisition of images can be modeled using Gaussian white noise. This type of noise is primarily caused by the Johnson-Nyquist noise (thermal noise). This type of noise can be modeled as an additive Gaussian noise that is independent at each pixel, adding a random value at each pixel.

A commonly used metric to measure the amount of noise existent in an image is the peak signal-to-noise ratio (PSNR). The PSNR is defined as the ratio of the mean squared error and the peak signal value of the image in log scale. The PSNR can be expressed as shown below.

$$PSNR = 10 \cdot \log_{10}(\frac{MAX^2}{MSE}),$$

where MAX is the max possible value of the image and MSE is the pixel-by-pixel mean squared error between the noisy image and the noiseless image.

3. Modeling Methodology

3.1. Simulation Configuration

The simulation first creates the interferometric images defined in the previous section. The simulation results without noise are created for differing heights and noiseless versions are stored. 0 mean Gaussian white noise with differing variation is added to the noiseless image. The newly created image and its PSNR value is stored.

The CNN model used for this work is a simple fully connected network. The combination of convolutional layers and maxpooling layers is used to create the final output. The input is a gray scale (single channel) surface plasmons resonance microscopy image of the interferometric pattern, and the output is the predicted height z given the input image.

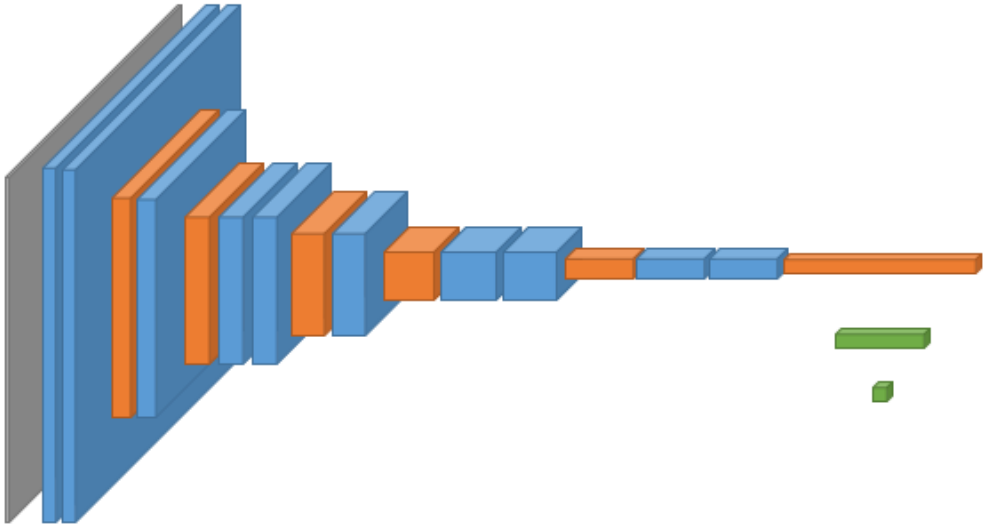

Figure 3.1.1. The convolutional neural network used for this work. Blue layers are convolutional layers, and orange layers are the maxpool layers. The gray layer is the input layer which takes the crop of a single interferometric pattern from the surface plasmon resonance microscopy, and the green layer is the fully connected layers used for regression. The model is used for regression rather than the more commonly seen classification.

The CNN network is trained using noisy images. This is to train the network to recognize real images that are used for this work. The level of noise for the real images is difficult to predict. Therefore, the amount of noise in the training data set is not fixed and is set contain a large range.

The training data set is augmented with random rotation, scaling, and cropping to account for real-world scenarios. The training data set is separated into the training and test data set. The model is trained until the test loss starts increasing. The model at that point is used for the analysis of interferometric images to prevent overfitting.

The model uses the rectified linear unit (ReLU) activation functions. Weight decay is used to prevent overfitting. The usage of dropout for regularization was found to be unreliable because of the regression nature of the given task of predicting the height of a given

interferometric image. During testing, the model was found to be unstable depending on which nodes were ignored. The usage of dropout was found to be not useful.

The input image is all rescaled such that the minimum pixel brightness is 0 and the maximum brightness is 1. This allows for the model to focus on the pattern of the interferometric image rather than look at certain points. The output target heights are also normalized to a [-1, +1] range. This normalization process improves the performance of the model significantly.

3.2. Optical Configuration

The optical setup to create and measure interferometric plasmonic microscopy images is shown in Fig. 3.2.1. A laser that radiates 633nm electromagnetic waves is used as the light source. The light is modulated such that the incoming light covers the entire field of view from the microscope while being in the TM mode.

The optical configuration used for the physical experiment is shown in Fig. 3.2.1.

A linear polarizer and a half-wave plate are used to change the polarization of the input laser into the TM mode. The piezo mirror is used to change the incident angle of the electromagnetic wave. A EMCCD camera is used to measure the light intensity at the output.

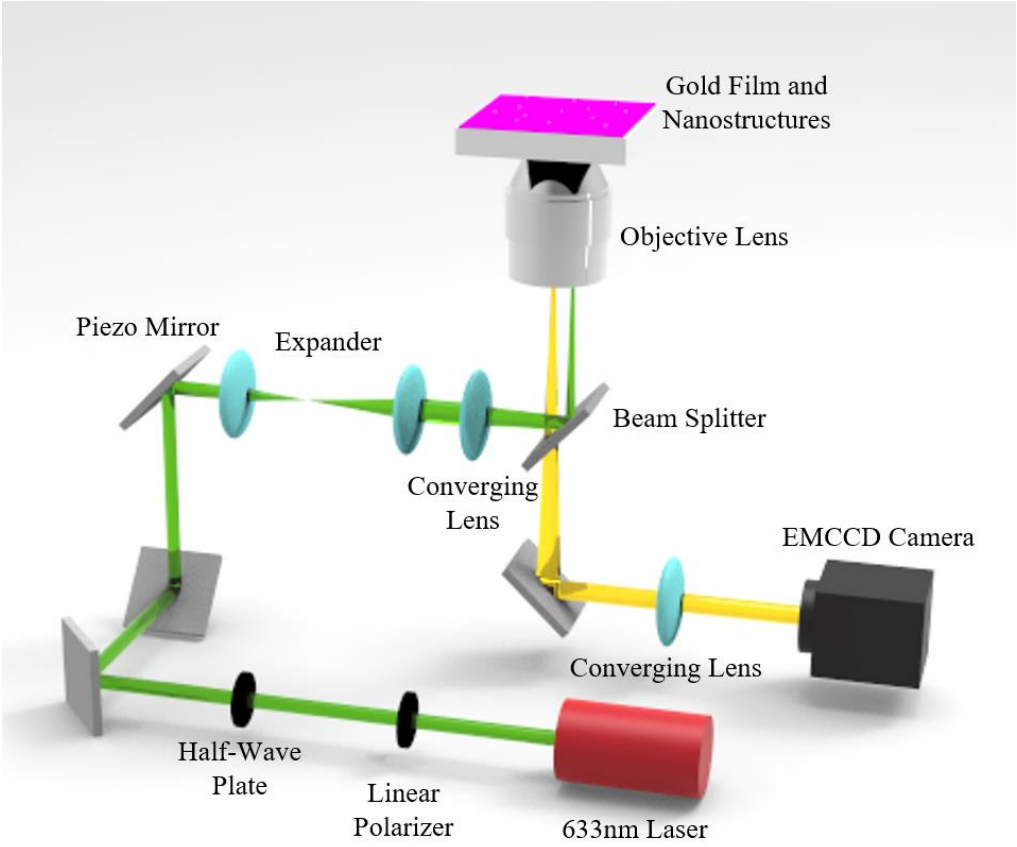

Figure 3.2.1. Schematics of the optical setup. The light source is a 633nm laser (red cylinder). The light is converted to TM mode using a linear polarizer and a half-wave plate. The light is expanded and focused on the sample on top of the objective lens. The output is focused using a converging lens and received by a EMCCD camera.

3.3. Sample Preparation

Two different types of samples are used for this work. A gold film with gold nanoparticles and a E-beam resist based nano-post array is created.

3.3.1. Gold Nanoparticle Films

The creation of gold nanoparticle films is first started by creating gold films. The gold film will be deposited onto a Bk7 glass slide. The sample preparation starts with cleaning the Bk7

slide. The Bk7 glass is cleaned using sonication. The sample is first cleaned with acetone for 3 minutes using sonication. This process is repeated for IPA then distilled water. The cleaned sample is then covered with gold using an evaporator. A 50nm thick film of gold is applied to the Bk7 surface. The addition of gold nanoparticles is done by utilizing a gold nanoparticle solution. A gold nanoparticle solution of 100nm diameters with an optical density (OD) of 1 suspended in a 0.1mM PBS solution is used for this work. The solution is diluted to a tenth of its original concentration is dropped on the gold coated Bk7 sample. The evaporation of water leaves just gold nanoparticles on the surface of the gold film.

3.3.2. E-Beam Resist Nano-post Array

The interferometric patterns of gold nanoparticles are very faint and hard to discern. The small nature of the nanoparticles makes it difficult to obtain clear interferometric patterns. E-beam resists can be used to easily obtain large nanostructures and create clear interferometric patterns.

The creation of E-beam resist structures starts in a relatively similar fashion to the creation of the gold nanoparticle film. The Bk7 film is first cleaned using the method used for gold nanoparticle films. The usage of an evaporator to coat the Bk7 sample is also identical. Afterwards, the gold film is coated with E-beam resist AR-N 7520 at 4000rpm angular velocity to create a coating 400nm thick. The E-beat resist is processed using E-beam lithography where the areas that are meant to be posts are exposed to E-beams. The sample is developed in a process that removes the parts of the E-beam resist not exposed to E-beams. The final sample is a gold film with a E-beam nano-post array.

3.4. Data Extraction

The method of obtaining interferometric images is to remove the background from the detected fields [15]. There are two methods of obtaining the original background, the spatial method and the temporal method.

The temporal method can be utilized by first obtaining an image of a clean sample. This means that the first taken image only contains the film and there are no aberrations on the surface.

The second image is taken after nanoparticles are dropped onto the surface of the sample. The difference between these two images is the interferometric pattern caused by the nanoparticles located on the surface of the film.

The spatial method can be utilized by obtaining an image from a location without any features of interest (nanoparticles or nanostructures). The sample can be moved such that two images are taken where one image does not have aberrations within the field of view and one image contains the nanoparticles and therefore the interferometric patterns. The difference of the two images is the interferometric caused by the nanoparticle.

Nanostructures require fabrication processes to be used on the sample before the sample can be observed using the microscopy setup. There is an inherent difficulty in creating nanostructures while it is being observed by a camera. Therefore, the temporal method is not viable to observe nanostructures. The spatial method is used to observe the nanostructures.

Nanoparticles are easier to observe using the temporal method. Readily available nanoparticles can be required in a solution. A solution with an appropriate density of the solution can be dropped onto the sample while it is being observed. Taking the first image with just the sample and then dropping nanoparticles on it to take the second image allows for the acquisition of interferometric images.

The taken images have a large field of view with multiple particles or structures. The images are processed by taking a crop of the pattern created by each structure. The cropped images are used for the final testing process.

3.5. Noise Implementation

The introduction of noise into the system to simulate real-world scenarios is done by adding a Gaussian white noise to the already created interferometric images. The interferometric images with different levels of noise are in Table 3.5.1.

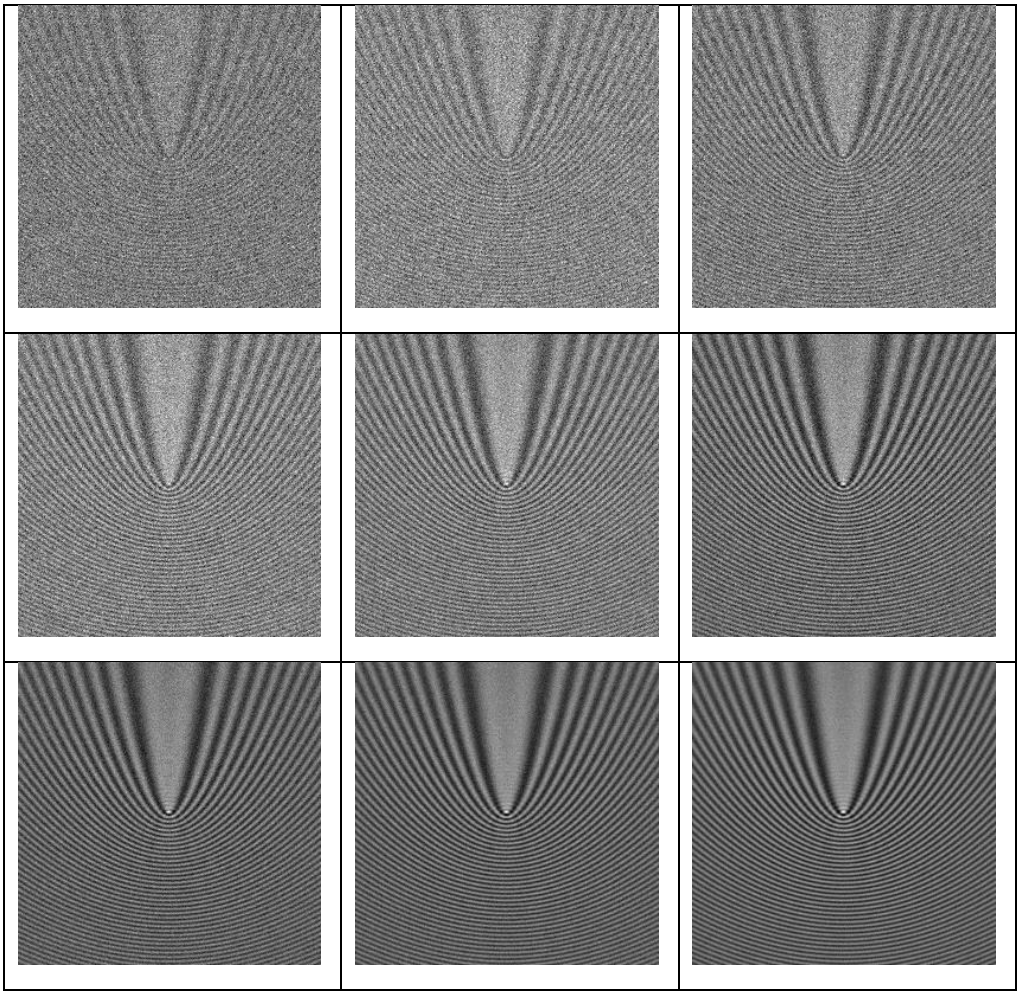

Table 3.5.1. The interferometric images with differing amounts of noise. Gaussian white random noise is added to the interferometric images to create the final noisy image.

4. Results

4.1. Simulation Results of the CNN Model

The MSE loss of the CNN model during training is in Fig. 4.1.1.

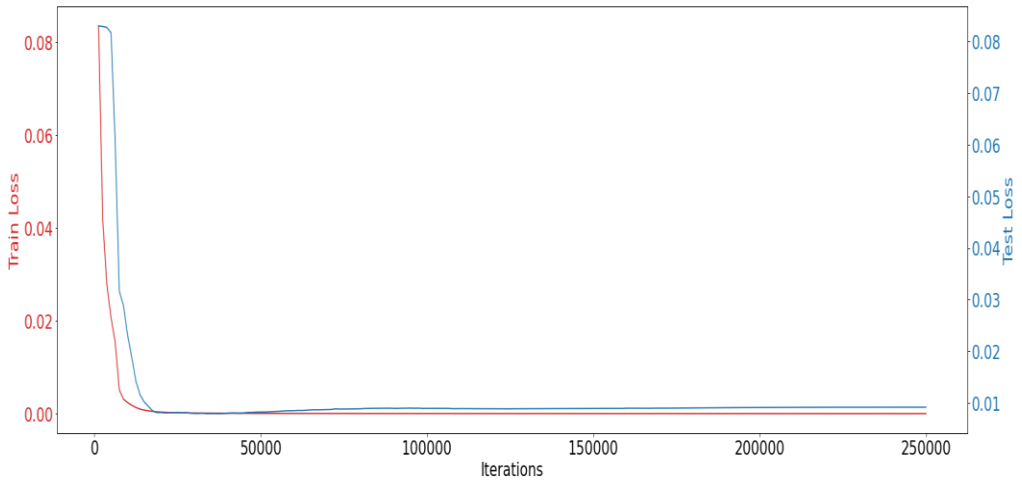

Figure 4.1.1. The training and test loss as the model is trained using simulated images. The training loss is constantly decreasing while the test loss is decreasing then increasing, implying that overfitting is occurring when training the CNN model.

The red line represents the training MSE loss, and the blue line represents the test MSE loss. The training and test loss are both decreasing as the number of iterations is increased. The model is able to accurately predict the z location of an interferometric image. The low MSE loss shown the test loss shows that the model is able to predict images that are not part of the training set. The ability of the model to predict new inputs accurately is proven.

4.1.1. Simulation Results Depending on the z Value

The model is only trained on noisy data to simulate real world data. The evaluation of the simulation is carried out by using the noiseless images. The predicted height of the particle and the real values are in Fig. 4.1.2.

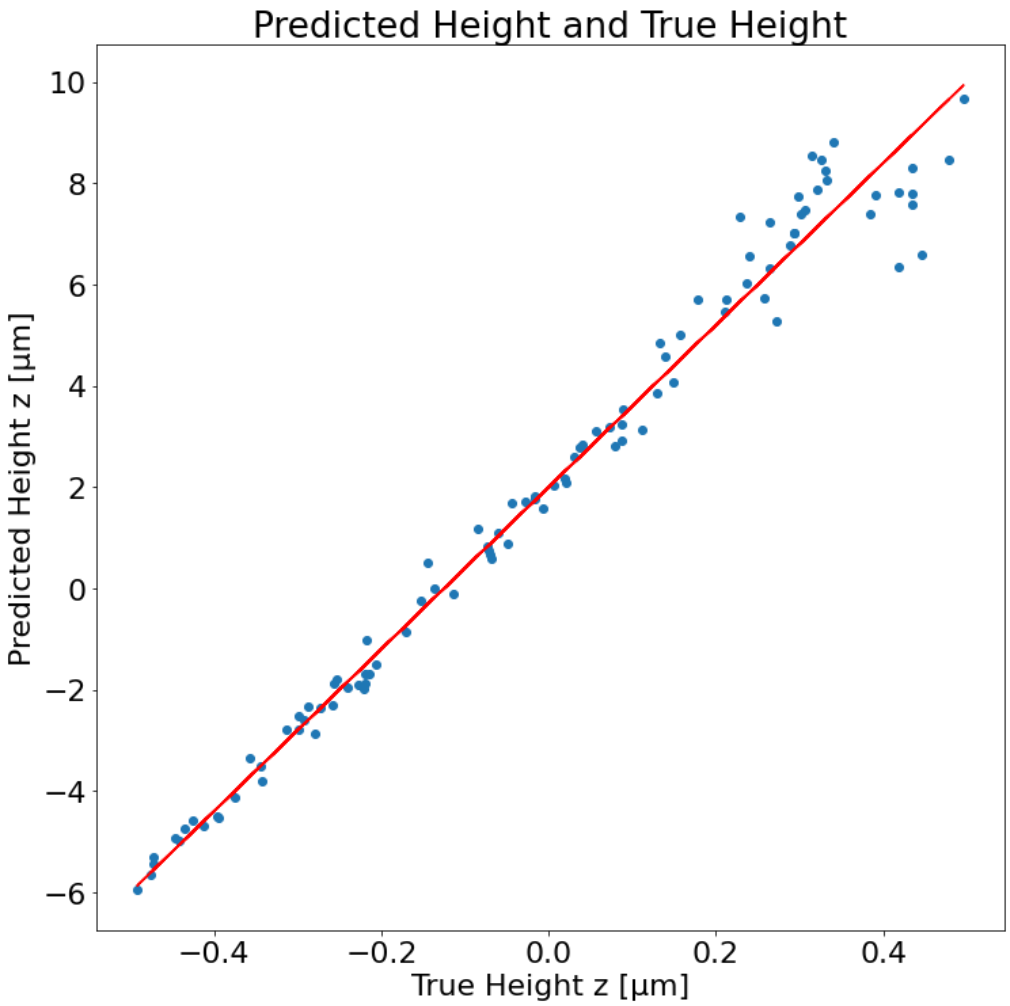

Figure 4.1.2. The prediction of heights given a noiseless image. The model shows a high accuracy when the back focal plane is near 0. The prediction becomes more difficult as the model moves further away.

The results show that as long as the z value is within a certain range, the model is able to predict the relative location to the back focal plane with high accuracy. This is especially true when the particle is near the back focal plane. The CNN model is able to accurately determine the location of the particle and therefore the height of the film given an interferometric pattern.

4.1.2. Simulation Results Depending on Noise

The result of testing given images withs different image noise levels is in Fig. 4.1.3.

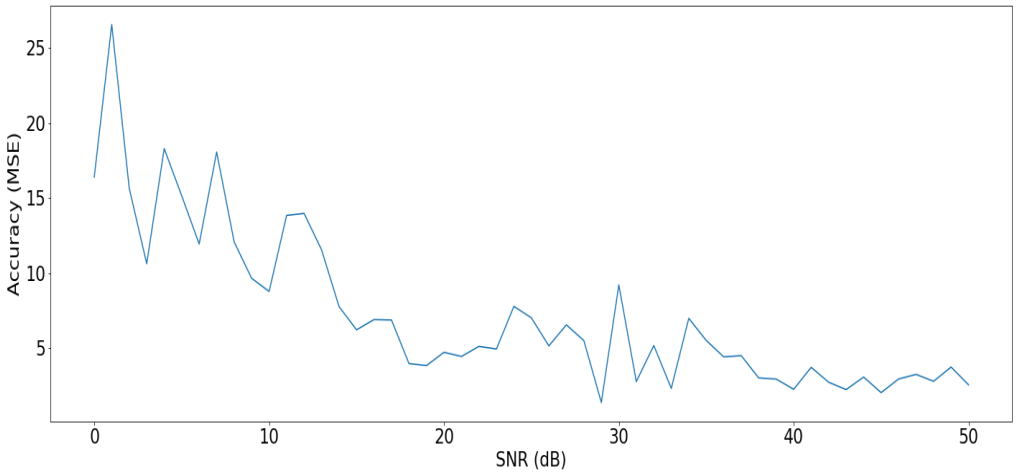

Figure 4.1.3. The prediction loss depending on the PSNR. The model prediction is closer to the ground truth when the noise level is low.

As expected, as the PSNR increases, the loss of the prediction decreases. The importance of an accurate imaging system with the ability to get images with less noise is highlighted.

4.2. Experimental Results with Gold Nanoparticles

The interferometric pattern created by the gold nanoparticles is shown in Fig. 4.2.1.

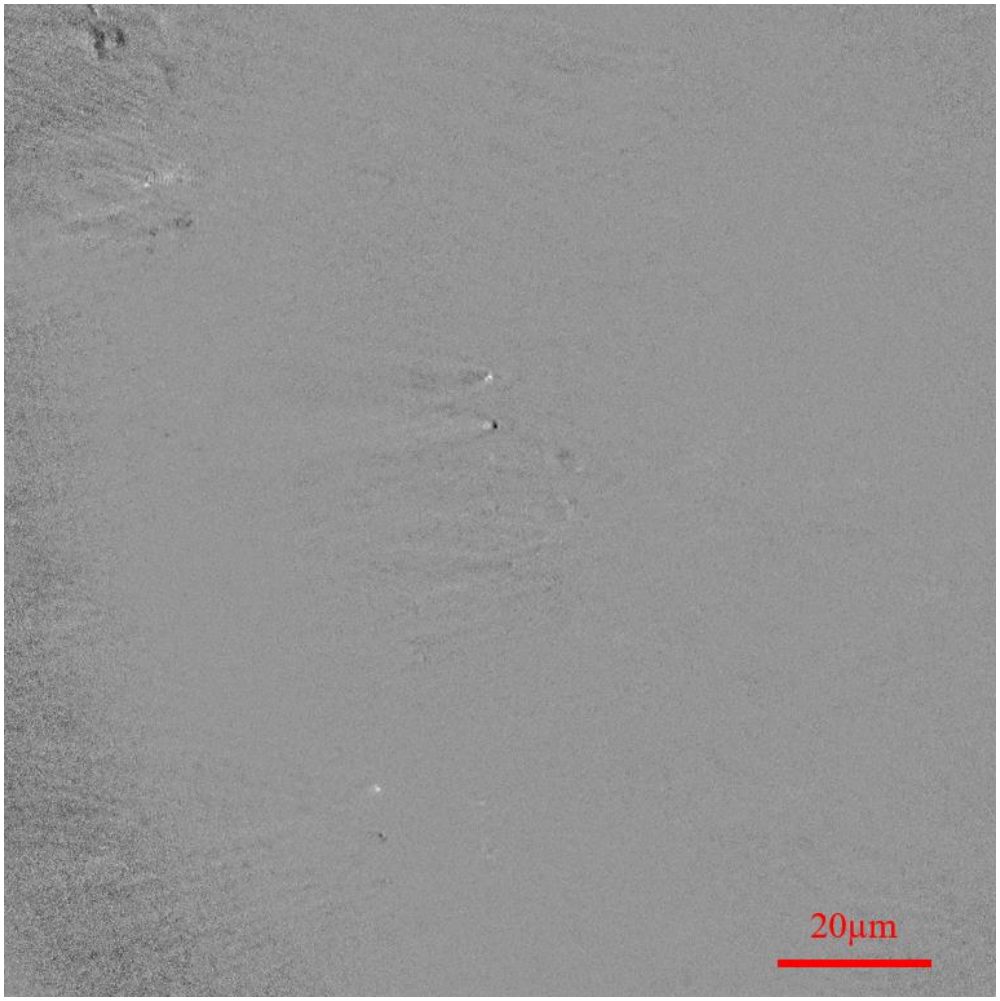

Figure 4.2.1. The interferometric pattern caused by 100nm gold nanoparticles. The vertical bright and dark patterns are complementary patterns created by the same particle. The small physical shift in the sample combined with the differential nature of the post-process image dictates that the final image contains complementary parts of the same pattern.

The spatial method is utilized to create the post-process images. The images are taken once at the location of gold nanoparticles. The sample is physically shifted, and another image is taken. The difference is the post-process image shown above. The gold nanoparticles are spread over a large area and the shifting the film by a large amount is difficult without the patterns caused by different gold nanoparticles to interfere with the final image. Therefore, the physical

shift of the sample is small. The difference between the two taken images creates a complementary inverse image that can be observed above as an inverted vertical shift.

The predicted heights for four interferometric patterns are shown in Fig. 4.2.2. The

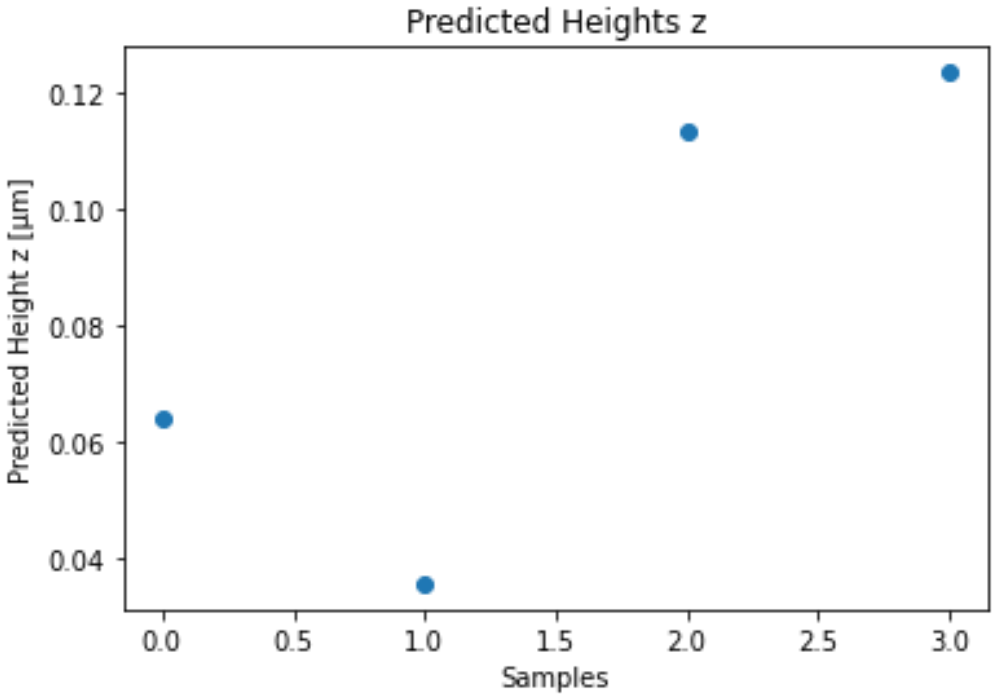

Figure 4.2.2. The predicted heights for the interferometric patterns. The gold nanoparticles are located on a film and the predicted heights should be close to each other. The predicted heights are very similar to each other.

The gold nanoparticles are located on a thin flat film. Their patterns are faint, implying that there might be complications with accurate predictions. However, the patterns created by gold nanoparticles create extremely similar values for the height when output from a trained neural network. This shows that the network is able to make nanoscale accurate results from even faint images.

4.3. Experimental Results with E-beam Resist Structures

The interferometric pattern created by an E-beam resist structure is shown in Fig. 4.3.1.

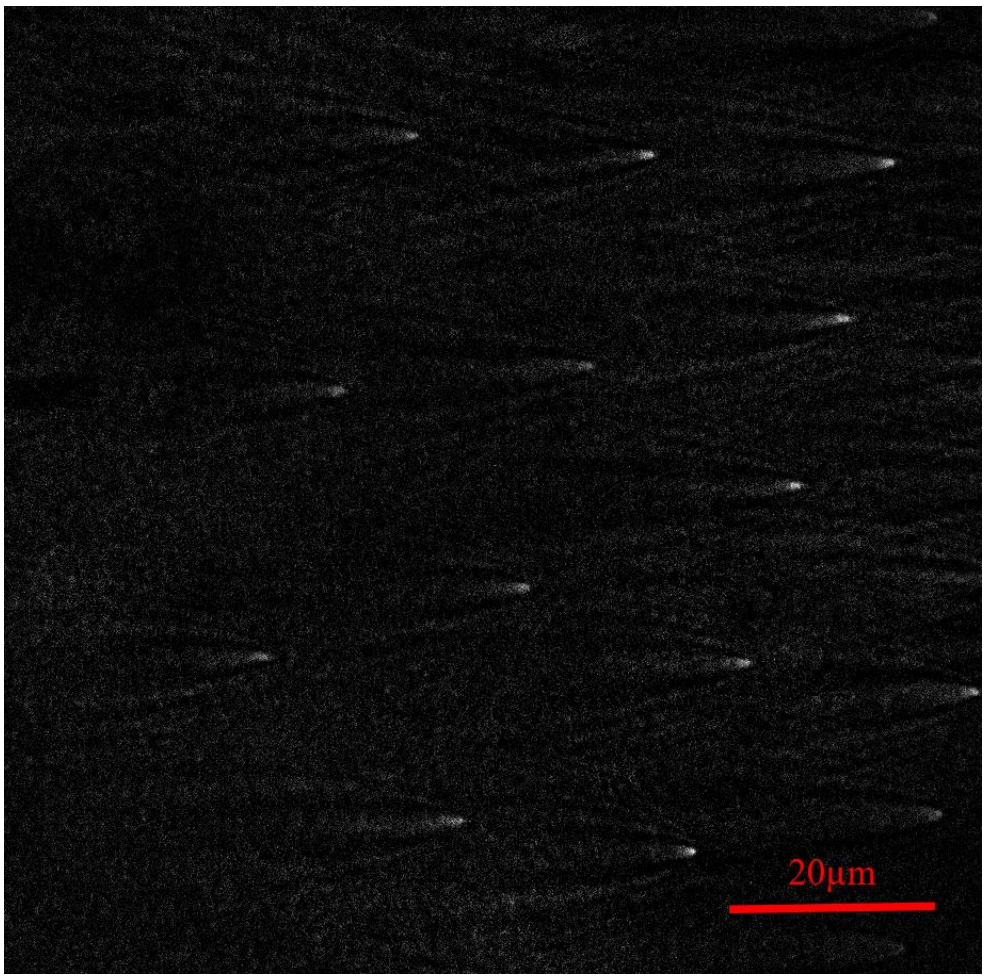

Figure 4.3.1. The interferometric pattern of a E-beam resist structure. The bright spots each correspond to a post structure. The interferometric pattern can be clearly seen for each post. The large post structure allows for a more distinct interferometric pattern. The large physical shift possible when using nanoparticles allows for the removal of complementary parts of interferometric patterns.

The image is taken by the spatial method. Two images are taken where one image is taken at

the location of the E-beam structures and one image is taken at the bare film. The difference between the two images is the post process image shown above. The interferometric patterns can be observed clearly above. E-beam structures are large compared to gold nanoparticles, and the created interferometric pattern is much more distinct.

The background image can be taken from a part of the sample completely free of any structures. Therefore, unlike the gold nanoparticle post-process image, the complementary double image does not exist, and a single interferometric image exists in the figure above.

The adaptation of the CNN model to the analysis of real images can be done by inputting the cropped images of the interferometric patterns. An example input to the neural network is shown in Fig. 4.3.2.

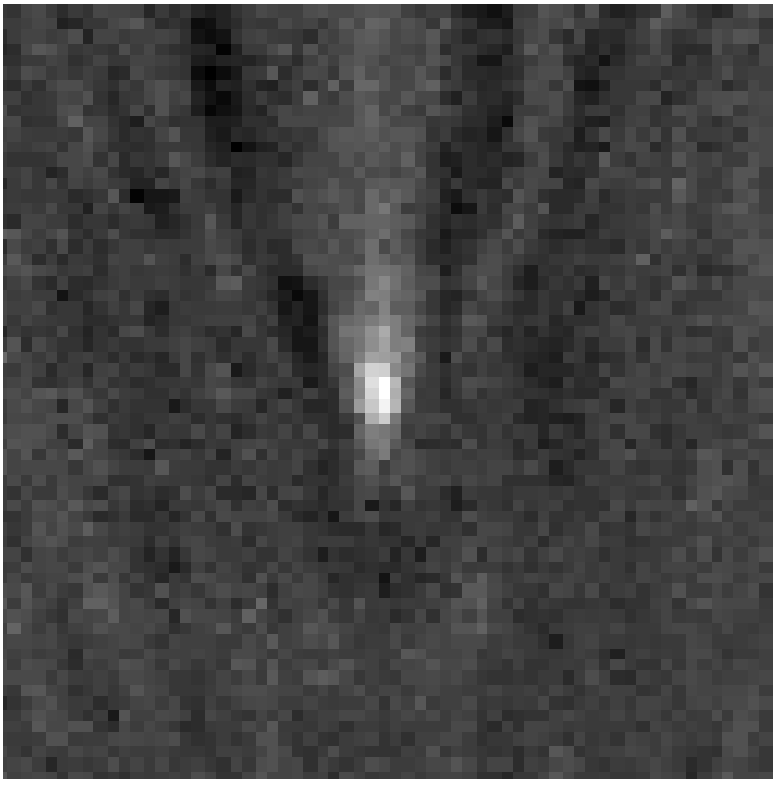

Figure 4.3.2. An example input to the created neural network. A cropped image centered on the center of the interferometric pattern is created and input into the neural network.

The result of testing the E-beam structure interferometric patterns is shown in Fig. 4.3.3.

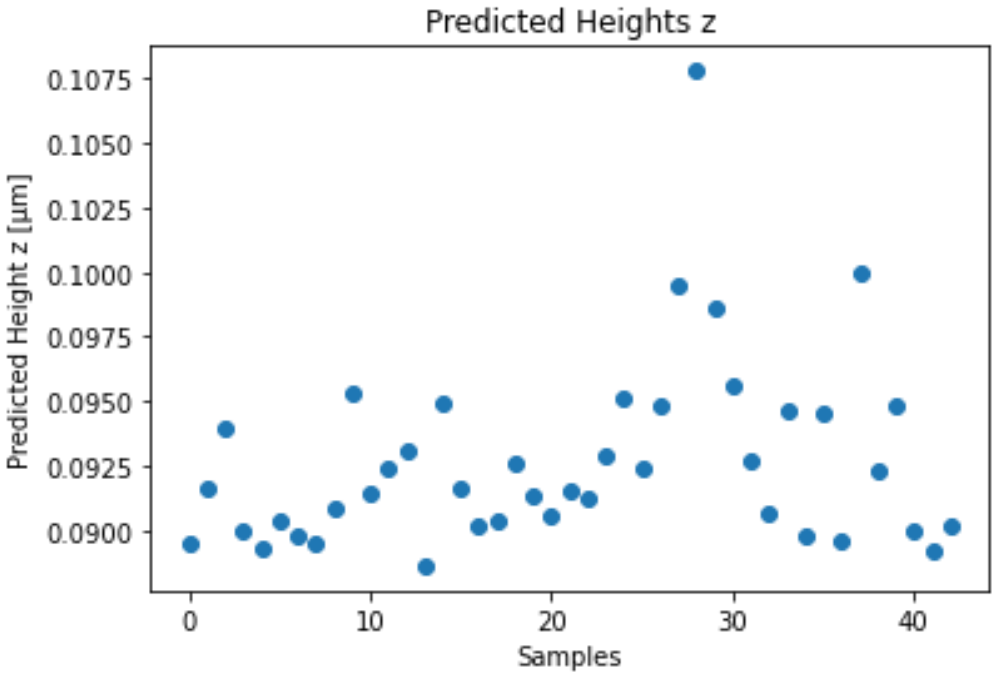

Figure 4.3.3. The output of the model given a E-beam resist structure interferometric image. Each point corresponds to a single pole structure. The posts are located on a flat film.

Figure 4.3.4. The atomic force microscopy (AFM) image of an E-beam resist nano-post array. The posts are created at a period of 19µm. The vertical lines shown in the image above are artifacts created by the oscillating cantilever during the creation of the AFM image.

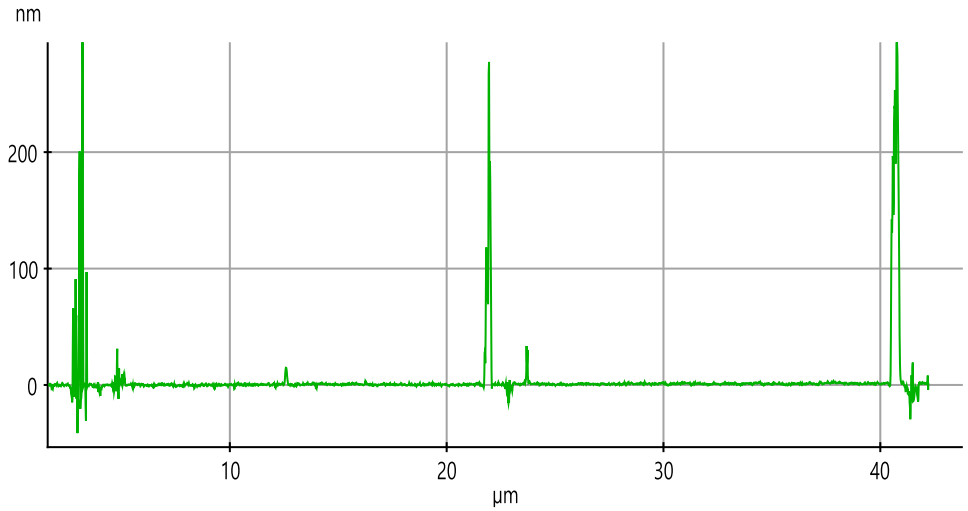

Figure 4.3.5. The horizontal slice of the nano-post array deposited on a gold film. The posts have similar heights on the gold film with a periodicity of 19µm.

The E-beam resist posts are located on a singular flat film as shown in the ground truth image created by the AFM image. The model is able to predict the heights of the posts within a small nanoscale range. The model is able to predict the relative locations of E-beam resist posts with nanometer accuracy. The interferometric imaging system is able to successfully add a new dimension to the imaging system with no modifications to the original surface plasmon resonance microscopy system.

5. Discussion

An overly complex model causes overfitting to the training data regardless of the number of epochs used to train the model. Nested CV is used to find the optimal configuration of the model. The model was found to work best when the number of parameters is somewhat limited. The task given for this work is simple and the given training data set is small. A more difficult task with a more extensive data set should require a more complex model to create accurate results.

An interesting observation made during this research is the usefulness of transfer learning for optimizing models created from only simulation data. The model is able to recognize and

properly estimate the output when only using simulated data. This holds true even in instances where the input image is particularly noisy. The simulation data is enough to create a deep learning model that can be used immediately with real world data to make accurate predictions. It has been observed that a lower test loss also leads to a lower error when using real images, implying that a strong correlation exists between theoretical simulated images and images taken from physical implementations that can be utilized by deep learning networks.

The network is in essence, learning a point spread function of interferometric plasmonic microscopy. The ability to analyze any interferometric pattern with a trained network further increases the generalizability and applicability of interferometric plasmonic microscopy.

6. Conclusion

This work aims to utilize deep learning techniques to retrieve more information from interferometric plasmonic microscopy systems. The deep learning model has been successful in retrieving the normal component from interferometric images with nanometer accuracy. The usage of simulated data to predict real-world phenomena with high accuracy is proven in this work but fine-tuning to easily attainable ground truth data should increase the performance of the model even further.

The ability to localize the nanostructures on a gold film using deep learning techniques also opens up the possibility of a similar method to track a nanoparticle with changing height relative to the film surface. The creation of 3D images with little modification to a traditional surface plasmon microscopy imaging system can be anticipated. The neural network used for this model is relatively simple, and a real-time system to acquire 3D images may be realizable.

The interferometric pattern seems to form a relation with the material of the nanostructure, especially their complex refractive index. Extensive analysis of different materials should allow for the creation of a data set with the interferometric pattern of different materials. It can be predicted that analysis of the interferometric pattern may allow for the extrapolation of information concerning not just the height of the nanostructure but also the material of the nanostructure.

Reference

[1] Zeng, Shuwen, et al. "Nanomaterials enhanced surface plasmon resonance for biological and chemical sensing applications." Chemical Society Reviews 43.10 (2014): 3426-3452.

[2] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).

[3] Polman, Albert, and Harry A. Atwater. "Plasmonics: optics at the nanoscale." Materials Today 8.1 (2005): 56.

[4] Raether, Heinz. "Surface plasmons on smooth surfaces." Surface plasmons on smooth and rough surfaces and on gratings (1988): 4-39.

[5] Poole Jr, Charles P. Encyclopedic dictionary of condensed matter physics. Academic Press, 2004.

[6] Maier, Stefan A. Plasmonics: fundamentals and applications. Vol. 1. New York: springer, 2007.

[7] Amendola, Vincenzo, et al. "Surface plasmon resonance in gold nanoparticles: a review." Journal of Physics: Condensed Matter 29.20 (2017): 203002.

[8] Zhou, Xiao-Li, et al. "Surface plasmon resonance microscopy: from single-molecule sensing to single-cell imaging." Angewandte Chemie International Edition 59.5 (2020): 1776-1785.

[9] Rothenhäusler, Benno, and Wolfgang Knoll. "Surface–plasmon microscopy." Nature 332.6165 (1988): 615-617.

[10] Son, Taehwang, et al. "Superlocalized three-dimensional live imaging of mitochondrial dynamics in neurons using plasmonic nanohole arrays." ACS nano 13.3 (2019): 3063-3074.

[11] Son, Taehwang, et al. "Surface plasmon microscopy by spatial light switching for label-free imaging with enhanced resolution." Optics letters 43.4 (2018): 959-962.

[12] Yu, Hui, et al. "Molecular scale origin of surface plasmon resonance biosensors." Analytical chemistry 86.18 (2014): 8992-8997.

[13] Yang, Yuting, et al. "Quantitative amplitude and phase imaging with interferometric plasmonic microscopy." Acs Nano 13.11 (2019): 13595-13601.

[14] Syal, Karan, et al. "Antimicrobial susceptibility test with plasmonic imaging and tracking of single bacterial motions on nanometer scale." ACS nano 10.1 (2016): 845-852.

[15] Yu, Hui, et al. "Plasmonic imaging and detection of single DNA molecules." ACS nano 8.4 (2014): 3427-3433.

[16] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[17] Shan, Xiaonan, et al. "Detection of charges and molecules with self-assembled nano-oscillators." Nano letters 14.7 (2014): 4151-4157.

[18] Magonov, Sergei N., and Myung-Hwan Whangbo. Surface Analysis with STM and AFM. VCH, 1996.

국 문 요 약

디포커스 간섭 플라즈몬 현미경 이미지를 딥러닝에 기반한 방법으로 분석 및 표면 프로파일링

표면 플라즈몬 현미경은 다용도 영상 기술이다. 샘플 표면 근처에 있는 현상을 선택적으로 촬영하는 것이 가능하며 간섭 패턴을 통해 일반 현미경에 비해 더 좋은 성능을 가진다. 샘플 필름 표면 위에 있는 나노 입자는 간섭 패턴을 형성하고 이는 필름에 대한 다양한 정보를 가지고 있다. 하지만 표면 플라즈몬 현미경을 통해 생긴 간섭 패턴은 회절한계보다 굉장히 크고 포물선 형태를 지니고 있기 때문에 일반적인 방법으로 분석하는 것이 어렵다. 본 연구에서는 딥러닝 기술을 이용해 표면 플라즈몬 현미경을 통해 생성한 이미지를 분석한다. 일반적인 convolutional neural network이 높은 정확도로 간섭 패턴을 가지고 있는 정보를 얻을 수 있었다. 딥러닝에 기반한 분석 방법으로 표면 프로파일링을 통해 이 기술의 가능성을 증명했다.

핵심되는 말: 머신 러닝, 딥러닝, 표면 플라스몬, 플라스몬 현미경, 디포커스 이미징, 표면 프로파일링