

本科毕业设计（论文）开题报告

专业：计算机科学与技术

班级：10011707

| | | | | | |
|----------|---|----|------|---------|-----|
| 学号 | 2017302430 | 姓名 | 许程远 | 指导教师 | 李学龙 |
| 报告题目 | 基于模态学习的光谱迁移方法 | | | | |
| 题目来源（划√） | 科研 <input checked="" type="checkbox"/> 生产 <input type="checkbox"/> 实验室 <input type="checkbox"/> 专题研究 <input type="checkbox"/> 企业联合 <input type="checkbox"/> | | | | |
| 论文类型（划√） | 工程设计类 <input type="checkbox"/> 实验研究类 <input checked="" type="checkbox"/> 软件开发类 <input type="checkbox"/> 其 他 <input type="checkbox"/> | | | | |
| 报告日期 | 2021 年 03 月 11 日 | | 报告地点 | 毅字楼 604 | |

1. 选题背景、意义及依据

随着计算机视觉的发展以及在大数据的驱动下，深度学习方法在音、视频领域涌现出各类应用，如图像分类、目标检测、图像分割、图像生成等。但这些应用大多是针对一种模态进行的学习，光谱迁移属于对跨模态数据间关联的研究。可见光图像可以呈现与人眼所见相似的信息，但对光照变化比较敏感，而红外图像可提供物体热量信息，在特定条件下，红外图像可看到可见光看不到的场景，从而延伸人类视觉感知范围。但红外图像又有数据量不足，细节缺失严重，获取代价大等缺点。因此，研究可见光图像转红外图像具有重大意义。本文将研究可见光光谱与红外光谱之间的对应关系，提出一种新的可见光图像转红外图像的方法。



(a)可见光图像 (b)红外图像

图 1. 可见光图像转红外图像示例，数据来源[1]

研究可见光光谱转红外光谱具有重大意义。光谱迁移目的在于探究光与光间的变换规律。目前的红外成像，往往存在细节缺失严重的问题，利用可见光图像带有的物体细节信息，可补充红外图像细节损失，提高红外成像质量。在军事领域，以红外形式呈现图像，可隐藏重要信息，降低信息泄露的风险，实现“视觉可见到视觉不可见”。红外图像目前数据量不足，获取困难，红外武器设计研制过程中若采用相关机器学习、深度学习方法，实现追踪、定位等任务，则不可避免需要大量数

据集做支撑，当探索清楚红外光谱与可见光光谱之间的转换关系后，可为相关军事任务拓展出全新的研发技术手段。除了可见光光谱和红外光谱之间迁移之外，我们未来还可以用类似方法探究紫外光，X 射线谱与可见光光谱之间的联系。

深度学习能够建立人工智能应用中遇到的多种数据概率分布模型，比如自然图像、包含语音的音频波形，以及自然语言语料库中的符号等。在深度学习开始阶段，最显著的成功在于判别模型的学习，通常是那些将高维、丰富的感官输入映射到类别标签的模型[2,3]。而对于生成模型的学习，传统方法拟合样本分布过程中，用到的最大似然估计、马尔可夫链等方法，由于存在难以近似的概率计算问题，以及难以在生成过程中利用深度学习分段线性单元优势等问题，发展较为缓慢。之后，对抗生成网络的出现[4]，解决了上述问题，为生成模型提供了一个新的框架。该框架中生成模型通过一个多层感知器传递随机噪声来生成样本，判别模型也是一个多层感知器。将该框架应用于图像翻译任务，可以解决该领域各种图像生成问题[5,6]。在图像翻译中，图像到图像转换(Image-to-Image Translation)是指从一个领域获取图像并将其转换为具有另一个领域图像的风格(或特征)的任务。例如，输入语义标签生成图像，输入黑白图像生成彩色图像，输入白天图像生成夜间图像。它们本质上都属于一种数据迁移过程。类似地，我们也可将此方法应用于可见光与红外光谱之间的数据迁移。

图像到图像转换任务往往涉及庞大的计算量。近些年来光学计算的发展给深度学习带来新的展望，基于光本身的干涉、衍射特点、3D 打印技术以及集成光子学的发展，使得打造全光机器学习平台成为可能，以光速执行机器学习任务，或采用光学硬件替换神经网络中的部分结构达到提高计算速度的目的，甚至光学芯片也可能在未来实现。并且，仿真技术的发展让我们可以通过计算机低成本模拟物理光学系统实验过程，可以预先判断实验原理可行性，从而指导光学平台搭建。

本文拟在探究基于光学计算方法对光谱迁移任务速度的提升和能耗的降低任务。

2. 国内外研究现状

目前学界主要将光谱迁移方法看作一类图像到图像转换任务，利用生成对抗网络的优势解决此问题。生成对抗网络因其在学习高维、复杂的真实数据分布方面的

潜能而受到机器学习领域的广泛关注。具体来说，它们不依赖于任何关于分布的假设，并且可以以一种简单的方式从潜在空间中生成真实样例[7]。这种强大的特性使得生成对抗网络可广泛应用，如图像翻译[8,9]、图像超分辨率生成[10]、视频生成[11]、图像转文字[12]等学术领域。

但生成对抗网络应用目前很多都受制于能耗和计算效率，基于硅基的电子元件虽然现在还能支撑，但随着时间推移，未来将难以满足不断增长的算力需求。若把光学系统看作传递信息的系统，它的作用和电子系统在本质上是相同的，在电子系统中，时间性传递随时间变化的电流和电压信号，而光系统中空间性传递复振幅或强度的空间分布。基于这样的共同本质，光学和电子信息传递领域之间具有紧密的联系。光相较于电子有着更快的传输速度，因而采用光学计算方法相较于电子计算方法的计算速度有着多个数量级的提升。

本文结合生成对抗网络强大的数据迁移能力和光神经网络加快计算速度的优势，提出一种新的可见光图像转红外图像方法。下面将分别对生成对抗网络和光神经网络的国内外研究现状进行介绍。

2.1. 基于对抗生成网络的图像翻译方法

为了适应不同图像翻译应用场景，可以通过施加不同损失函数约束实现。除了传统对抗生成网络 Sigmoid 交叉熵损失之外，文献[13]中损失基于红外图像和可见光图像本身的性质加入了亮度特征，使得生成图像的亮度更接近于真实图像。文献[14]中的损失函数中加入了鉴别器网络中的参数，即比较生成图像与真实图像的鉴别器网络参数差异性，在生成过程中增加约束，使得生成图像更接近真实图像。其他方法还包括条件随机场[15]、SSIM 度量[16]，特征匹配[17]，非参数损失[18]，基于匹配协方差统计量的损失[19]等。虽然上述各方法对特定任务都有良好的损失度量效果，但面对不同的场景，需要重新设计损失函数，可移植性较差，即没有一个面向各类图像到图像翻译任务统一的框架。

GAN 其中的生成器 (Generator) 和判别器 (Discriminator) 也可根据应用面不同进行不同的替换。文献[20,21]中采用多对生成器，分阶段生成图像，但计算量较大。[9,13]中的生成器参考医学图像分割常用的 U-Net 网络[22]，包括一个捕获上下文信息的收缩路径和一个允许精确定位的对称拓展路径，往往是针对小数据集完成端到

端的训练,有利于浅层特征与深度语义信息融合。至于辨别器现在使用最广泛的是 PatchGAN[23],属于马尔可夫辨别器,分块检测输入图像,在图像到图像翻译任务中对细节纹理信息保留较好。马尔可夫随机场和深度神经网络的结合[24]使得图像合成能够分别保留局部特征和全局特征。例如文献[8,25]中无监督双向对抗生成网络,文献[9,13]有监督对抗生成网络的辨别器部分都使用 PatchGAN。

综上,我们推测生成器采用 U-Net,辨别器采用 PatchGAN 应当是较为合理的,在文献[9]中也已经证明该组合是图像翻译任务典型的选择。除了以上关于网络结构的优势,文献[9]还为图像翻译任务提供了统一的框架,不需要根据任务类型的不同而设计特定的损失函数,迁移性强。因此,我们的红外光谱迁移网络主框架受文献[9]的启发设计。

2.2. 基于深度学习的光神经网络框架

现有的光神经网络分为两个框架:

一种是利用透射层(或反射层)上每个点之间的空间光衍射来模拟神经元之间的信息传输[26,27,28,29]。通过傅里叶光学[30]标量衍射理论和光学成像系统的频率特性给我们的启发,无论是相干成像系统还是非相干成像系统都可以看作线性系统,而人工神经网络中的连接层都是由线性构成,二者具有很大的相似性。并且透镜的位相调制和傅里叶变换性质告诉我们一个透镜也可以实现对输入图像的卷积操作,这与卷积神经网络中的卷积操作也有很大相似性。有很多将此种光学原理应用于深度神经网络的例子,如通过光-电转换实现混合光电卷积神经网络[29,31,32,33,34]。文献[31]中采用光学前端(卷积层)加传统神经网络(全连接层)后端实现图像分类,光学信息处理中的最典型的 4f(非)相干滤波系统可实现输入信息与掩膜(Mask)脉冲响应的卷积运算,利用该原理我们即可设计透镜阵列和合适的掩膜(可类比于卷积神经网络中的 kernel)并行的对输入图像进行卷积运算操作。文献[32]给我们提供了一种纯光学卷积神经网络计算系统,系统包括卷积层,池化层,非线性激活层,全连接层,每一层的计算都是由光学完成,只有不同设备之间数据的传递通过电子传输完成。并且,不同于其他的光电卷积神经网络,这种系统易于扩展,即可以通过对相关设备的复用来增加卷积层数。首先卷积层利用空间光调制器(SLMs)和傅里叶透镜阵列实现,下采样(池化层)采用光学步进卷积

(optical-strided convolution)实现, 非线性采用 sCOMS 相机曲线实现, 全连接层采用光学点积实现。文献[33]通过正负光电探测器的使用改良了全光衍射深度神经网络对光强非负性的限制。

另一种框架由光学干涉单元和光学非线性单元构成, 如相干纳米光子电路和光子神经网络[35,36,37], 该框架得益于集成光子学的发展。该框架可应用于其他大量使用矩阵乘法和非线性激活的人工神经网络算法, 包括卷积神经网络和递归神经网络。如文献[35]中芯片中集成的光神经网络单元主要分为两个部分: 光学干涉单元(OIU)以及光学非线性单元(ONU)。前者实现矩阵乘法, 后者实现非线性激活。而[36]中线性层采用空间光调制器(SLM)和傅里叶透镜实现, 而非线性激活函数是在具有电磁感应透明度的激光冷却原子中实现。文献[37]介绍了一种光神经网络中非线性激活函数的电光硬件平台。光-光非线性通过将一小部分输入光信号转换成模拟电信号来工作, 模拟电信号用于对原始光信号进行强度调制, 而不降低处理速度。但此框架对集成光子学设备要求较高, 成本较高, 对于 CNNs 的实现而言, 使用光学 4f 系统比使用声光调制器阵列建立光学卷积单元结构是更好的实现卷积的方法。

本文从光学结构简单、易于实现、仿真有效性、对硬件光学平台搭建有指导性帮助的角度出发, 选择第一种框架来研究光-电混合卷积神经网络实现可见光图像转红外图像的课题。

2.3. 现存问题

现有关于对抗生成网络实现图像翻译的研究主要关注于如下问题 1)生成对抗网络训练过程中, 可能无法收敛到理想值, 生成对抗网络在训练中要求生成器生成目标域图像, 不一定会让判别器同时达到最佳判别模型, 二者在博弈过程中可能会不断的抵消对方的进步; 2)生成器梯度消失问题, 当判别器非常准确时损失很快收敛到零, 从而无法提供可靠的路径使生成器的梯度继续更新, 造成生成器梯度消失。3)模型坍塌的问题, 生成器坍塌到了一个极狭小的分布内, 生成的样本不在变化, 即生成器在某种情况下重复生成完全一致的图像。

光-电混合深度神经网络的研究主要关注于如下问题 1)光-电混合神经网络中大量数据在光领域和电子领域间来回转换, 其中的能耗和延迟不可忽略, 存在光-电转换成本较高的问题。2)低功率光学非线性难以实现, 光学非线性相对较弱, 实现

往往需要大的相互作用长度和高信号功率的组合。

3. 课题研究目标、研究内容、研究方法及关键技术

3.1. 研究目标

将光学神经网络引入图像翻译任务，其主要的目的在于：

- 1) 通过将生成器中的部分卷积层替换为光学前端，提高生成图像速度。
- 2) 降低生成对抗网络中的电子部分训练参数，提高训练速度。
- 3) 处理大数据时，有效的降低能源消耗。

3.2. 研究内容

本毕业设计拟在使用光-电卷积对抗神经网络实现图像翻译任务的研究。

- 1) 物理实验中需要纳米光子学超表面硬件实现光学掩膜，存在成本高且不易调试的问题。研究利用计算机快速傅立叶变换(FFT)算法和角谱传播器模拟衍射等方法计算各阶段光场分布，从仿真实验的角度，缓解成本高且不易调试的问题。
- 2) 多卷积层光学替换的光-电信号转换过程中，可能会产生高延迟问题。研究输入信号为光信号条件下第一卷积层光学替换方法，从减少光-电信号转换次数，简化网络框架角度，缓解延迟问题。
- 3) 可见光图像转红外图像过程中，可能存在生成图像效果不好度量的问题。研究 PNSR, SSIM 度量方法对红外图像生成结果进行评价。从图像信噪比，结构相似性的量化角度解决生成图像效果不好度量的问题

3.3. 研究方法

受对抗生成网络结构[9]和光-电混合卷积神经网络结构[31,34]的启发，在"U-Net"网络中引入光学 4f 相关器卷积结构，提出如下图 2 网络结构设计。

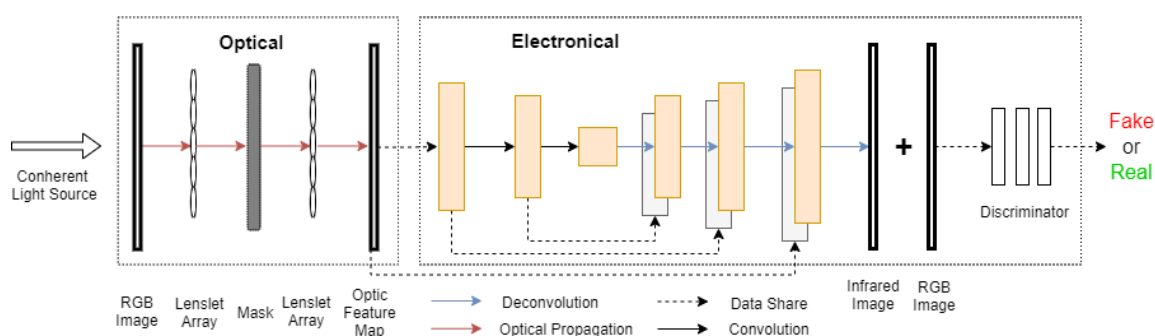


图 2. 红外光谱迁移网络结构图

生成器部分是改良的关键，我们将第一个卷积层替换为光学卷积层，即 $4f$ 相关器(上图左下部分)。在菲涅耳近似下，由于透镜在其前后焦平面的光场分布之间提供了精确的傅立叶变换关系，串联的两个透镜在输入处提供了光场的符号反转傅立叶变换。而正是这样的傅里叶变换关系，在相干光学线性系统中即可实现卷积操作，原理如下：

均匀平面光照在图像上，由于图像不同位置具有不同的相位、振幅透过率，因此，通过图像的光便携带了相应的图像信息，之后光通过第一个透镜，根据[30]所述透镜的傅里叶变换性质，空间域的图像信息在掩膜(Mask)面前被转换为频域信息，即掩膜平面前的光信息代表了图像在频域的特征。通过设置掩膜的透光率和厚度，对入射光分别进行振幅和相位的调制。通过掩膜后，得到图像频域信息与掩膜频域调制函数的乘积关系，我们都知道在频域的乘积相当于在空间域的卷积，这样就能实现相应的卷积操作。光再经过一倍焦距的自由空间的传播，相当于进行了傅里叶逆变换，图像频域的光信息被还原为图像空间域光信息，只不过此时已经是与掩膜发生卷积后的结果。最后光再通过最后一个透镜，球面波被还原为平面波，使得我们可以在任意距离进行观察。

由于 CNN 需要很多卷积操作，本文网络架构由一个 $4f$ 相关器阵列(Lenslet Array)组成，以使单个层的所有卷积操作能够并行发生。透镜阵列需要具有足够小的孔径，使阵列中的每个部分光传播都可以被视为傍轴，使得透镜的傅里叶变换性质有效。[34]中也已经证明各个 $4f$ 相关器之间无串扰，使我们可以把每个 $4f$ 相关器看作一个独立的卷积单元。我们可以通过代码仿真此光学部分，利用快速傅立叶变换(FFT)算法和角谱传播器模拟衍射方法，计算不同传播距离下的光场分布。

完成了卷积后的光信息通过非线性平方定律光电探测器，将卷积结果光信号转换为电信号，传入计算机中进行后续的神经网络操作。对抗生成网络中鉴别器(Discriminator)参考[23]中 Patch-GAN 的鉴别器部分进行设计。

最后，关于图像生成效果度量一直是一个开放而又困难的问题。文献[38]中采用类似图灵测试的方法，通过人工标注，对图像翻译效果进行评价，存在易受人为主观因素影响，定性不定量的问题。传统的度量标准，如逐像素均方误差(perpixel mean-squared error)，不评估结果的联合统计，因此对整体结构没有良好的考量。综

合考虑,我们采用结构相似性(SSIM)以及峰值信噪比(PNSR)的定量度量方法对红外生成图像进行评估。

4. 论文所遇到的困难和问题、拟采取的解决措施及预期达到的目标

4.1. 困难和问题、拟采取的解决措施

- 1) 相关研究较少。目前光-电神经网络方面的研究大多停留在图像分类任务上,将此引入图像翻译任务中是一个大胆的尝试。
- 2) 涉及信息在光、电两种模态间的转换,其转换规律尚不明确,研究面临巨大挑战。可以利用深度学习强大的非线性表达能力,基于傅里叶光学原理利用 4f 相关器实现快速光学卷积,设计端到端的转换网络,避免从原理不通的窘境。
- 3) 深度学习模型需要大量数据支撑。OTCBVS 数据集中大部分没有可见光图像和红外图像成对数据,不能进行有效监督训练。FLIR 红外目标识别数据集中图像对没有对准。初步选择 KAIST 数据集进行训练,后期可采用特殊感光 CCD 元件直接收集红外图像,对模型进行训练。

4.2. 预期达到的目标

- 1) 提出一种新的可见光图像转红外图像方法。之前的方法中没有将光学原理应用于图像到图像翻译任务的尝试。
- 2) 在不影响图像翻译评价指标峰值信噪比(PSNR (Peak Signal-to-Noise Ratio))、结构相似性(SSIM (Structural Similarity))、图像翻译效果直观评价指标 AMT(Amazon Mechanical Turk)的前提下,提高图像生成速度。
- 3) 撰写论文一篇。

5. 论文进度安排

2021 年 10 月 30 日-2021 年 01 月 29 日,了解毕业设计的内容和要求,确定题目,学习傅里叶光学、对抗生成网络基础知识;

2021 年 01 月 30 日-2021 年 02 月 15 日,深入研究相关文献资料,研究分析其使用方法及其优缺点;

2021 年 2 月 16 日-2021 年 03 月 1 日,初步方案的设计和确定;

2021 年 03 月 2 日-2021 年 03 月 12 日,完成并上传开题报告;

2021 年 03 月 13 日-2021 年 04 月 19 日, 初步模型搭建与调试;

2021 年 04 月 20 日-2021 年 04 月 26 日, 得出模型结果, 完成论文中期报告;

2021 年 04 月 27 日-2021 年 05 月 17 日, 完善模型, 评估模型, 分析结果, 撰写毕业论文;

2021 年 05 月 18 日-2021 年 05 月 30 日, 完成研究结果的整理, 上传毕业论文, 毕业设计答辩。

6. 参考文献

- [1] S. Hwang, J. Park, N. Kim, Y. Choi, and I. Kweon, "Multispectral Pedestrian Detection: Benchmark Dataset and Baseline.," in *CVPR*, 2015.
- [2] G. Hinton, L. Deng, D. Yu, G. Dahl, et al, "Deep Neural Networks for Acoustic Modeling in Speech Recognition ," in *IEEE Signal processing magazine*, 2012.
- [3] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, pp. 1097-1105, 2012.
- [4] I. Goodfellow, J. Abadie, M. Mirza, B. Xu, et al, "Generative Adversarial Nets," in *arXiv preprint arXiv*, 2014.
- [5] M. Ravanbakhsh, M. Nabi, E. Sangineto, L. Marcenaro, et al, "Abnormal Event Detection in Videos Using Generative Adversarial Nets," in *IEEE International Conference on Image Processing*, pp. 1577-1581, 2017.
- [6] Y. Chen, K. Lai, J. Liu, "CartoonGAN: Generative Adversarial Networks for Photo Cartoonization," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9465-9474, 2018.
- [7] J. Hong, U. Hwang, J. Yoo, S. Yoon, "How Generative Adversarial Networks and Their Variants Work: An Overview, " in *ACM Computing Surveys (CSUR)*, pp.1-43, 2019.
- [8] J. Zhu, T. Park, P. Isola and A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,"in *IEEE International Conference on Computer Vision*, pp. 2242-2251, 2017.



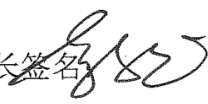
- [9] P. Isola, J. Zhu, T. Zhou and A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks, " in *CVPR*, pp. 5967–5976, 2017.
- [10] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, et al, "Photorealistic Single Image Super-Resolution Using a Generative Adversarial Network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.
- [11] C. Vondrick, H. Pirsiavash, and A. Torralba, "Generating Videos With Scene Dynamics," in *Advances in Neural Information Processing Systems*, pp. 613–621, 2016.
- [12] K. Lin, Q. Li, D. He, Y. Zhang, and T. Sun, "Adversarial Ranking for Language Generation," in *The Thirty-First Annual Conference on Neural Information Processing Systems*, 2017.
- [13] X. Qian, M. Zhang and F. Zhang, "Sparse GANs for Thermal Infrared Image Generation from Optical Image," in *IEEE Access*, vol. 8, pp. 180124-180132, 2020.
- [14] C. Wang, C. Xu, C. Wang and D. Tao, "Perceptual Adversarial Networks for Image-to-Image Transformation," in *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4066-4079, 2018.
- [15] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected," in *International Conference on Learning Representations*, 2015
- [16] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image Quality Assessment: from Error Visibility to Structural Similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [17] Dosovitskiy and T. Brox, "Generating Images with Perceptual Similarity Metrics Based on Deep Networks," in *Conference and Workshop on Neural Information Processing Systems*, 2016.
- [18] C. Li and M. Wand, "Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis," in *IEEE Conference on Computer Vision and*

Pattern Recognition, 2016.

- [19] J. Johnson, A. Alahi, and F. Li, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," in *European Conference on Computer Vision*, 2016.
- [20] H. Zhang, T. Xu, S. Li, et al. "StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks," in *Proceedings of the IEEE international conference on computer vision*, pp. 5907-5915, 2017.
- [21] Y. Liu, O. Tuzel, "Coupled Generative Adversarial Networks," in *arXiv preprint arXiv*, 2016.
- [22] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation, " in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, 2015.
- [23] C. Li and M. Wand, "Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, " in *European Conference on Computer Vision*, 2016.
- [24] C. Li and M. Wand, "Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2479-2486, 2016.
- [25] Z. Yi, H. Zhang, P. Tan and M. Gong, "DualGAN: Unsupervised Dual Learning for Image-to-Image Translation," in *2017 IEEE International Conference on Computer Vision*, pp. 2868-2876, 2017.
- [26] X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All Optical Machine Learning Using Diffractive Deep Neural Networks," in *Science*, 2018.
- [27] K. Zhou, L. Fang, T. Yan, M. Wu, P. Li, T. Fan, Q. Wu, X. Lin, and H. Dai, "In Situ Optical Backpropagation Training of Diffractive Optical Neural Networks," in *Photonics Research*, vol. 8, issue 6, pp. 940-953, 2020.
- [28] K. Dou, Y. Deng, T. Yan, Q. Wu, X. Lin, and H. Dai, "Residual D²NN: Training Diffractive Deep Neural Networks via Learnable Light Shortcuts," in *Optics Letters*,

vol. 45, issue 10, pp. 2688-2691, 2020.

- [29] D. Mengu, Y. Luo, Y. Rivenson and A. Ozcan, "Analysis of Diffractive Optical Neural Networks and Their Integration with Electronic Neural Networks," in *Journal Selection Topics in Quantum Electron*, vol. 26, no. 1, pp. 1-14, 2019.
- [30] J. Goodman, Introduction to Fourier Optics, 2008.
- [31] J. Chang, V. Sitzmann, X. Dun, W. Heidrich and G. Wetzstein, "Hybrid Optical-Electronic Convolutional Neural Networks with Optimized Diffractive Optics for Image Classification," in *Scientific Reports*, vol. 8, no. 1, pp. 12324, 2018.
- [32] Y. Gu, S. Gao, and Z. Liu, "Optronic Convolutional Neural Networks of Multi-Layers with Different Functions Executed in Optics for Image Classification," in *Optics Express*, vol. 29, no. 4, pp. 5877-5889, 2021.
- [33] X. Li, D. Mengu, Y. Luo, Y. Rivenson, A. Ozcan, "Class-Specific Differential Detection in Diffractive Optical Neural Networks Improves Inference Accuracy," in *Advanced Photonics*, vol. 1, no. 4, pp. 46001, 2019.
- [34] S. Colburn, Y. Chu, E. Shlizerman and A. Majumdar, "Optical Frontend for a Convolutional Neural Network," in *Applied Optics*, vol. 58, issue 12, pp. 3179-3186, 2019.
- [35] C. Shen, N. Harris, S. Skirlo, et al, "Deep Learning with Coherent Nanophotonic Circuits," in *Nature Photon*, 2017.
- [36] Y. Zuo, H. Li, J. Zhao, Y. Jiang, Y. Chen, P. Chen, G. Jo, W. Liu, and W. Du, "All-Optical Neural Network with Nonlinear Activation Functions," in *Optica*, vol. 6, issue 9, pp. 1132-1137, 2019.
- [37] Williamson, T. Hughes, M. Minkov, B. Bartlett, S. Pai, H. Fan, "Reprogrammable Electro-Optic Nonlinear Activation Functions for Optical Neural Networks," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, pp. 1-12, 2020.
- [38] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, et al, "Improved Techniques for

| | |
|--|--|
| Training Gans, " in <i>NIPS</i> , 2016. | |
| <p>指导教师意见:</p> <p>同意开题。</p> <p style="text-align: right;">  签名 </p> <p style="text-align: right;">2021 年 3 月 12 日</p> | |
| <p>开题评议小组成员: </p> | |
| <p>开题评议小组意见: (包括对论文的选题、难度、进度、工作量、论文形式意见):</p> <p>1. 论文选题: <input type="checkbox"/> 有理论意义; <input type="checkbox"/> 有实用价值; <input checked="" type="checkbox"/> 有理论意义与实用价值; <input type="checkbox"/> 意义不大。</p> <p>2. 论文的难度: <input type="checkbox"/> 偏高; <input checked="" type="checkbox"/> 适当; <input type="checkbox"/> 偏低。</p> <p>3. 论文的工作量: <input type="checkbox"/> 偏大; <input checked="" type="checkbox"/> 适当; <input type="checkbox"/> 偏小。</p> <p>4. 进度: <input checked="" type="checkbox"/> 可行; <input type="checkbox"/> 不可行;</p> <p>5. 学生开题报告中反映出的综合能力和表达能力: <input checked="" type="checkbox"/> 好; <input type="checkbox"/> 较好; 一般; <input type="checkbox"/> 较差。</p> <p>6. 论文形式意见: <input checked="" type="checkbox"/> 可行; <input type="checkbox"/> 不可行;</p> <p>7. 对论文选题报告的总体评价: <input checked="" type="checkbox"/> 好; <input type="checkbox"/> 较好; <input type="checkbox"/> 一般; <input type="checkbox"/> 较差。</p> <p>(在相应的方块内作记号“√”)</p> | |
| <p>评 议 结 论</p> | <p>是否同意论文选题报告: <input checked="" type="checkbox"/> 同意; <input type="checkbox"/> 需重做 (在相应的方块内作记号“√”)</p> <p style="text-align: right;">  评议小组组长签名 </p> <p style="text-align: right;">2021 年 3 月 12 日</p> |