資工專題競賽 學生大放異彩

資訊工程專題是本校資工系重要的必修課程之一，同學們自由報名參加專題競賽，經由同儕間互相觀摩與切磋，展現資工系多元豐富研究成果，同時也強化了資工系同學們的凝聚力。以下是得獎作品介紹：

特優

專題題目：Designing Area-Efficient Ray-Triangle Intersection Hardware Unit in GPU RT-Core

學生姓名：甯宇綸

指導教授：葉宗泰教授

專題介紹：

這個專案專注於 GPU RT 核心中的資源消耗型光線 - 三角形交集單元。我們將 Moller-Trumbore 交集演算法拆解為三個階段。通過篩選式硬體設計，每個階段都排除可能不相交的三角形，以降低後續階段對硬體的需求，從而減少整體面積。我們的方法在 FPGA 合成中實現了面積上的優勢，且性能損失在可接受範圍內。我們還設計了符合 AXI4 協定的交集單元，並將其部分整合進系統，包括 Xilinx 的 AXI4 快取，並使用 Genesis ZU FPGA 平台進行驗證。

優等

專題題目：Research on Predictive Networks Based on Multivariate Short Time Series Inputs and Transformer Architecture

學生姓名：賴怡暄、吳宜靜、楊芊華

指導教授：黃敬群教授

專題介紹：

隨著時間序列預測技術的進步，基於 Transformer 的模型在處理長時間序列預測任務中展現了令人印象深刻的性能。然而，這些模型在短期序列預測中的表現仍有待提升。這主要是因為模型架構在有效捕捉短序列特徵方面存在一定限制，從而影響了預測的準確性。為了提高基於 Transformer 的模型在短期序列預測中的表現，我們對最先進的 Crossformer 模型進行了以下優化： 首先，我們引入了 ProbSparse Self-attention，這一改進精煉了原有的 Crossformer 路由器機制。通過在計算 Attention Scores 之前對 Query 矩陣進行初步篩選，我們保留了重要的 Query 值，使模型能夠基於顯著特徵做出更準確的預測。 其次，我們引入了混合專家模型（MoE），以處理輸入數據的多樣化分布和時間序列的不同特徵。MoE 利用不同的前饋網路（FFNs）專家來處理序列中的不同特徵，並適應多變量數據，使信息能夠由多個 FFNs 同時處理，從而提升模型的適應性。 這些改進使得模型能夠更有效地處理短期序列中的局部特徵，從而提高了在小型數據集（ILI）和大型數據集（ETTh1）上的預測準確性。

優等

專題題目：Learning Diffusion Models with Occlusion Handling for Facial Landmark Detection

英文姓名：曾家祐、侯博軒

指導教授：林彥宇教授

專題介紹：

人臉特徵點偵測（Facial Landmark Detection, FLD）的目的是檢測人臉影像上的特定關鍵點。準確的人臉關鍵點檢測對於許多應用至關重要，包括人臉識別、表情分析和虛擬實境等。在我們的研究中，通過將處理期望的面部關鍵點的擴散模型（diffusion models）進行條件化。我們充分利用擴散模型在處理噪聲樣本方面的優勢，從而恢復正確的關鍵點，同時訓練我們的模型能夠識別從錯誤關鍵點到正確人臉關鍵點位置的反向過程。擴散模型在恢復噪聲樣本方面表現出色。因此，我們提議探索其在處理當前 FLD 方法中因人臉影像遮蔽引起的檢測誤差的可能性。我們的模型具有改進錯誤人臉關鍵點的能力，最終可以作為所有 FLD 方法的精煉器。

佳作

專題題目：Video Change Detection via Transformer-Based Architecture

學生姓名：文玠敦、馬楷翔、莊書杰

指導教授：陳冠文教授

專題介紹：

我們的專案探討了使用基於 Transformer 的模型 TransCD 來檢測影片中的變化。與傳統方法需要在時間和空間維度上都進行對齊不同，我們的方法僅對齊時間維度，從而在光線不足和視角極端變化等高難度條件下提升了性能。我們實施了自適應幀搜尋技術，以動態對齊幀，並使用微調技術來調整模型權重以提高準確性。我們的結果顯示，這種方法不僅提高了效率，同時保持了高品質的檢測效果，展示了 Transformer 架構在推進影片分析技術的潛力。

佳作

專題題目：Designing Fixed-Point Transcendental ISAs for Heterogeneous TinyML Acceleration

學生姓名：蔡承翰、范均宏

指導教授：葉宗泰教授

專題介紹：

這個專案設計了定點數先驗指令集架構（ISA），旨在提升 RISC-V CPU 上的 TinyML 性能，主要聚焦於 MobileViT 模型。通過將複雜操作分解為基本組件，並實作專門的 ISA 和設計對應的硬體單元，我們實現了顯著的加速。我們使用 CFU-Playground 平台在 NEXYS A7-100T FPGA 板上進行測試，發現在 TensorFlow Lite for Microcontrollers（TFLM）核心中，我們的自定義指令使得 MobileViT 模型推理速度比基準 CPU 快了 1.56 倍。相比於傳統的查找表（LUT）方法，我們的方法提供了更大的靈活性和記憶體效率，支持多種激活函數而無需硬體修改。

佳作

專題題目：Video Deblurring and Interpolation with Motion-Aware Transformer

學生姓名：朱致伶

指導教授：林彥宇教授

專題介紹：

我們提出了一個創新的運動感知的 Transformer 模型，用於同時進行視頻去模糊和插幀的雙重任務。該模型通過從多個模糊圖像中恢復高幀率且清晰的視頻，有效解決了運動模糊問題。運動模糊是由於曝光過程中物體的持續運動所造成的。為了解決這個問題，我們的運動感知 Transformer 模型充分利用了時間資訊，通過在視頻中多個連續模糊幀間共享 Intra-Motion 和 Inter-Motion 提示來實現。這些運動提示記錄了像素運動的大小和方向。Intra-Motion 提示捕捉單個模糊幀內的像素運動，而 Inter-Motion 提示則捕捉相鄰模糊幀間的像素運動。我們使用類似 UNet 的運動元素提取器 (motion extractor) 來預測這些運動提示，並將其輸入到視頻去模糊和插幀的 Transformer 中，從而降低了這些任務的複雜性並提升了模型性能。我們的研究表明，所提出的運動元素提取器顯著提升了視頻去模糊和插幀任務的效果。我們將持續改進運動提示在視頻去模糊和插幀 Transformer 模型中的整合與應用。

佳作

專題題目：BEVGaussian: Generate Scene-level 3D Gaussian from BEV image

學生姓名：李杰穎

指導教授：劉育綸教授

專題介紹：

在本研究中，我們的目標是從鳥瞰圖（BEV）圖像中生成場景級的 3D 高斯分佈，包括衛星圖像、 高度圖和語義圖。現有的方法僅限於物件等級的生成，無法產生場景級的三維高斯分佈。為了解決這個問題，我們利用現有的物件生成技術，從 BEV 圖像中創建高品質的 3D 高斯分佈，然後將這些物件整合到場景中。我們的方法不需要訓練，並且通過使用 BEV 圖像作為輸入，可以輕鬆地修改三維場景。

Students Excel in Computer Science Project Competition

The Computer Science project is a crucial mandatory course in NYCU's Computer Science Department. Students voluntarily enter project competitions, where they collaborate and observe each other's work, highlighting the department's diverse and substantial research achievements. This experience also fosters greater cohesion among Computer Science students. Below, we introduce the award-winning projects:

Excellence Award

Project Title: Designing Area-Efficient Ray-Triangle Intersection Hardware Unit in GPU RT-Core

Student: Yu-Lun, Ning

Advisor: Dr. Tsung-Tai Yeh

Project Introduction:

This project focuses on the resource-consuming ray-triangle intersection unit within the GPU RT-core. We broken down the Moller-Trumbore intersection algorithm into three stages. Through a sieve-like hardware design, each stage eliminates triangles that may not intersect to reduce the need of hardware in subsequent stages to decrease the overall area. Our method achieves an area advantage in FPGA synthesis, and the performance loss is acceptable. We have also designed the intersection unit to comply with the AXI4 protocol, partially integrating it into system includes Xilinx's AXI4 cache, and verified it using the Genesis ZU FPGA platform.

First-Class Award

Project Title: Research on Predictive Networks Based on Multivariate Short Time Series Inputs and Transformer Architecture

Students: Lai Yi-Xuan, Yang Chien-Hua, Wu Yi-Jing

Advisor: Dr. Ching-Chun Huang

Project Introduction:

With advancements in time series forecasting technology, Transformer-based models have demonstrated impressive performance in predicting tasks involving long time sequences. However, their performance in short-term sequence forecasting still falls short. This may be due to the model architecture's limitations in effectively capturing features of short sequences, thus restricting prediction accuracy. To enhance the performance of Transformer-based models in

short-term sequence forecasting, we optimize the state-of-the-art (SOTA) Crossformer model with the following improvements: Firstly, we introduce ProbSparse Self-attention, which refines the original Crossformer Router mechanism. By performing preliminary filtering on the Query matrix before calculating the Attention Scores, we retain only the important Query values, enabling the model to make more accurate predictions based on significant features. Secondly, we introduce Mixture of Experts (MoE) to handle the diverse distribution of input data and the varied characteristics of time series. MoE utilizes different Feed Forward Networks (FFNs) experts to process distinct features within the sequence while accommodating multivariate data, allowing information to be processed by multiple FFNs to improve the model's adaptability. These enhancements enable the model to more effectively handle local features in short-term sequences, thereby improving prediction accuracy on both small datasets (ILI) and large datasets (ETTh1).

First-Class Award

Project Title: Learning Diffusion Models with Occlusion Handling for Facial Landmark Detection

Students: Bosyuan Hou, Chiayu Tseng

Advisor: Dr. Yen-Yu Lin

Project Introduction:

Facial landmark detection(FLD), aims to detect specific key points on facial images. The accurate detection of facial landmarks is essential for many applications, including face recognition, expression analysis, and virtual reality.

In our work, we propose to adapt a new approach towards handling the task of FLD, by conditioning diffusion models on the desired facial landmark points. We utilize the strength of diffusion model learning to recover correct key points from noisy samples, and train our model to recognize the backward process from erroneous key points to correct facial landmark locations.

Diffusion models yield a strong performance on recovering noisy samples. Therefore, we propose to explore its potential on dealing with detection error in current FLD methods caused by facial image occlusion. Our model has the ability to refine erroneous facial landmarks, and eventually act as a refiner for all FLD methods.

Merit Award

Project Title: Video Change Detection via Transformer-Based Architecture

Students: Kai-Siang Ma, Shu-Chieh Chuang, Chieh-Dun Wen

Advisor: Dr. Kuan-Wen Chen

Project Introduction:

Our project explores the use of TransCD, a transformer-based model, to detect changes in videos. Unlike traditional methods that require alignment in both temporal and spatial dimensions, our approach aligns only the temporal dimension, enhancing performance under challenging conditions like poor lighting and extreme viewpoint changes. We implemented adaptive frame search to dynamically align frames and fine-tuning techniques to adjust model weights for better accuracy. Our results show that this method not only improves efficiency but also maintains high detection quality, demonstrating the potential of transformer architectures in advancing video analysis technology.

Merit Award

Project Title: Designing Fixed-Point Transcendental ISAs for Heterogeneous TinyML Acceleration

Students: Cheng-Han Tsai, Chun-Hong Fan

Advisor: Dr. Tsung-Tai Yeh

Project Introduction:

This project designs fixed-point transcendental ISAs for enhancing TinyML on RISC-V CPUs, focusing on the MobileViT model. By decomposing complex operations into basic components and implementing specialized ISAs and designing a corresponding hardware unit, we achieve significant acceleration. Using the CFU-Playground platform on NEXYS A7-100T FPGA boards, our custom instructions in TensorFlow Lite for Microcontrollers (TFLM) kernels result in MobileVit model inference a speedup of 1.56X over the baseline CPU. Our approach offers greater flexibility and memory efficiency compared to traditional Lookup Table (LUT) methods, supporting multiple activation functions without hardware modifications.

Merit Award

Project Title: Video Deblurring and Interpolation with Motion-Aware Transformer

Students: Chu Chih Ling

Advisor: Dr. Yen-Yu Lin

Project Introduction:

We propose a novel Motion-Aware Transformer model for the dual task of video deblurring and interpolation. This model addresses motion blur by recovering high-frame-rate, clear videos from multiple blurry images, achieving both video deblurring and frame interpolation simultaneously. Motion blur in images is caused by the continuous movement of objects during exposure. To address this, our Motion-Aware Transformer fully utilizes temporal information through Intra-Motion and Inter-Motion Prompts, shared among multiple consecutive blurry images in videos. The motion prompts store the magnitude and direction of pixel motion. The Intra-Motion Prompt captures pixel motion within a single blurry frame, while the Inter-Motion Prompt captures pixel motion between adjacent blurry frames. By predicting motion prompts with a UNet-like motion extractor and using these prompts as input to the video deblurring and interpolation transformer, we reduce the complexity of these tasks and improve model performance.

Our work demonstrates that the proposed motion extractor significantly enhances the performance of video deblurring and interpolation tasks. We are continuing to improve the blending and utilization of motion prompts in the video deblurring and interpolation transformer model.

Merit Award

Project Title: BEVGaussian: Generate Scene-level 3D Gaussian from BEV image

Students: Jie-Ying Lee

Advisor: Dr. Yu-Lun Liu

Project Introduction:

In this work, we aim to generate scene-level 3D Gaussians from bird's-eye view (BEV) images, including satellite images, heightmaps, and semantic maps. Existing methods are limited to object-level generation and cannot produce scene-level 3D Gaussians. To address this, we leverage existing object generation techniques to create high-quality 3D Gaussians from BEV images and then integrate these objects into the scene. Our approach is training-free and allows for easy modification of the 3D scenes by using BEV images as input.