วงศธร ภาธรสุวรรณ : การแยกแยะเสียงพยาธิวิทยาบนพื้นฐานโครงข่ายประสาทคอนโวลูชันแบบหลายมาตราส่วน (PATHOLOGICAL VOICE DETECTION BASED ON MULTI-SCALE CONVOLUTIONAL NEURAL NETWORK)
อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.พีระพงษ์ อุฑารสกุล และผู้ช่วยศาสตราจารย์ ดร.คมเดช ภาพัฒนบุรี, 111 หน้า

คำสำคัญ: การตรวจจับเสียงพยาธิวิทยา/โครงข่ายประสาทคอนโวลูชัน/ปัญญาประดิษฐ์

วิทยานิพนธ์นี้นำเสนอวิธีการใหม่ในการตรวจจับเสียงทางพยาธิวิทยาโดยใช้สถาปัตยกรรมโครงข่ายประสาทเทียมคอนโวลูชันแบบหลายมาตราส่วน (Multi-Scale Convolutional Neural Network: MSConvNet) วัตถุประสงค์ของการศึกษานี้คือการระบุความผิดปกติของเสียงจากข้อมูลเสียงที่ไม่ผ่านกระบวนการใด ๆ สถาปัตยกรรมใหม่นี้ถูกเรียกว่า RS-MSConvNet ถูกออกแบบมาเพื่อวินิจฉัยความผิดปกติของเสียงจากข้อมูลเสียงดิบ โมเดลนี้ใช้บล็อกคอนโวลูชันหลายมาตราส่วนเชื่อมต่อกับชั้นเชื่อมโยงสมบูรณ์ (Fully Connected Layer: FC) สำหรับการจำแนกประเภทโดยมุ่งหวังที่จะแยกแยะความแตกต่างของเสียงที่มีความผิดปกติและเสียงที่ปกติ

นอกจากนี้ วิทยานิพนธ์ฉบับนี้ยังนำเสนอ RS-MSConvNet-SVM ซึ่งเป็นโมเดลการเรียนรู้แบบผสมผสานระหว่างการเรียนรู้เชิงลึกและการเรียนรู้ของเครื่อง ซึ่งรวมความสามารถในการสกัดคุณลักษณะของ RS-MSConvNet กับประสิทธิภาพการจำแนกประเภทของ Support Vector Machine (SVM) เพื่อเพิ่มความแม่นยำในการระบุความผิดปกติของเสียง อีกทั้งยังใช้กลไกการคัดเลือกคุณลักษณะอย่าง Particle Swarm Optimization (PSO) ซึ่งเป็นวิธีการเพิ่มประสิทธิภาพ ผลการทดลองกับฐานข้อมูล TORGO ซึ่งประกอบด้วยตัวอย่างเสียงที่มีสุขภาพดีและเสียงที่ผิดปกติ ผลการทดสอบพบว่า RS-MSConvNet, RS-MSConvNet-SVM และ RS-MSConvNet-SVM ที่ใช้ PSO สามารถบรรลุความแม่นยำได้ 86.46%, 87.61% และ 88.09% ตามลำดับ ซึ่งผลลัพธ์แสดงให้เห็นว่าวิธีที่นำเสนอมีประโยชน์สำหรับการตรวจจับเสียงทางพยาธิวิทยา

สาขาวิชา วิศวกรรมโทรคมนาคม
ปีการศึกษา 2566

ลายมือชื่อนักศึกษา .....
ลายมือชื่ออาจารย์ที่ปรึกษา .....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

WONGSATHON PATHONSUWAN: PATHOLOGICAL VOICE DETECTION BASED ON MULTI-SCALE CONVOLUTIONAL NEURAL NETWORK. THESIS ADVISOR: ASSOC. PROF. PEERAPONG UTHANSAKUL, AND ASST. PROF. KHOMDET PHAPATANABURI, Ph.D. 111 PP.

Keyword: PATHOLOGICAL VOICE DETECTION/END-TO-END ARCHITECTURE /MULTI-SCALE CONVOLUTION/SPATIALTEMPORAL FEATURE/HYBRID MODEL

This thesis proposes a new method for identifying pathological voice patterns by utilizing Multi-Scale Convolutional Neural Network (MSConvNet) architectures. The aim of the study is to detect abnormal voice characteristics from unprocessed speech data. A new architecture, namely RS-MSConvNet, has been designed to detect abnormal voice from raw speech. This model uses a multi-scale convolution block, a spatiotemporal feature block, and a fully connected layer for classification, with the goal of capturing differences between abnormal voice and normal voice.

Furthermore, the thesis proposes the RS-MSConvNet-SVM, a hybrid model that combines the feature extraction capabilities of RS-MSConvNet with the classification power of Support Vector Machine (SVM) to improve the accuracy of speech pathology identification. In addition, it utilizes a feature selection mechanism that employs Particle Swarm Optimization (PSO), a computational technique that enhances performance. Thorough experimentation with the TORGO database, which includes both normal and abnormal speech samples, revealed that the RS-MSConvNet, RS-MSConvNet-SVM, and RS-MSConvNet-SVM with PSO achieved remarkable accuracies of 86.46%, 87.61%, and 88.09%, respectively. The outcomes show that our proposed methods are useful for pathological voice detection.

School of Telecommunication Engineering Student's Signature .....

Academic Year 2023 Advisor's Signature .....

Co-Advisor's Signature .....