

Título: Progressos para a melhoria do reconhecimento automático da fala de crianças

Nome: Thomas Rolland

Doutoramento em Engenharia Informática e de Computadores

Orientador (es) Alberto Abad Gareta

Resumo

Nos últimos anos, o reconhecimento automático da fala tem tido um avanço considerável, proporcionando novas oportunidades em diversas áreas de aplicação direcionadas a crianças. No entanto, a fala de criança apresenta desafios únicos, que têm um impacto negativo no desempenho dos sistemas de reconhecimento convencionais. Esta tese aborda estratégias para superar tais desafios, com foco na variabilidade acústica e na escassez de dados específicos.

Inicialmente, a investigação concentra-se no desenvolvimento de sistemas de reconhecimento híbridos, baseados em abordagens de transferência de conhecimento. É introduzido um novo método de aprendizagem multilingue, que combina aprendizagem multi-tarefa e transferência de conhecimento, mostrando-se superior em contextos com recursos limitados.

Posteriormente, são explorados os contributos dos diferentes componentes das arquiteturas baseadas em transformers para a adaptação para crianças. A nova abordagem de ajuste parcial destaca-se pela sua superioridade em relação ao ajuste global do modelo. O uso de módulos adaptadores também é investigado como método de transferência eficaz, demonstrando um bom desempenho em comparação com métodos convencionais.

Com base na eficácia dos adaptadores, é proposto o método de ajuste de adaptadores de via dupla, que utiliza dados de conversão de texto para fala para aumentar a disponibilidade de dados reais, resultando em melhorias consideráveis.

Por fim, são explorados métodos alternativos, culminando no novo conceito de adaptadores partilhados, nos quais um adaptador é usado em todas as camadas do modelo. Estes adaptadores oferecem uma eficiência de parâmetros superior, mantendo um bom desempenho de reconhecimento, e tornando-se numa escolha promissora para modelos de reconhecimento para crianças.

Em conclusão, esta tese proporciona uma visão abrangente e metodologias inovadoras para enfrentar os desafios do reconhecimento de fala de criança, contribuindo significativamente para o avanço nesse campo. Os resultados abrem caminho para uma melhor interação pessoa-máquina em aplicações educacionais, de entretenimento, saúde e tecnologia assistida focadas especificamente para crianças.

Palavras-chave: Reconhecimento automático da fala, fala de criança, transferência de conhecimento, texto para fala, eficiência de parâmetros