

RESUMO

BONILLA C, D A. *Reconhecimento Automático dos Fonemas do Língua Portuguesa Usando Modelos de Composição de Redes Neurais Especialistas*. 2016. 130f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2016.

Os sistemas de reconhecimento automático de fala para a detecção de fonemas proporcionam vantagens para o reconhecimento *online* de fala representada por um sinal de som. Os estudos na área de inteligência computacional têm permitido associar as vantagens do processamento paralelo de informação à distribuição da carga computacional visando simplificar o uso de modelos complexos de sistemas. O desenvolvimento de um sistema de reconhecimento automático de fala envolve vários processos que abrangem várias áreas da pesquisa, tais como a linguística, o processamento de sinais e a inteligência computacional. Nesta dissertação, o processo inicia-se com o pré-processamento do sinal de fala fornecido na entrada do sistema, visando extrair e representar de modo mais sucinto as características principais relacionadas ao sinal em um dado instante de tempo. Inspirado pelo preceito que recomenda “dividir para conquistar”, um modelo de composição de redes neurais especialistas é explorado, permitindo dividir o espaço de decisão do problema complexo de reconhecimento de fala para que cada especialista cuide somente de uma área bem delimitada deste espaço de decisão. Vale ressaltar que cada especialista incluído no modelo composto precisa tratar e tomar uma decisão a respeito de cada uma das amostras pré-processadas. O conjunto de decisões assim obtidas pelos especialistas são ponderadas. Desse modo, o sistema especialista que estiver com maior peso na saída acaba determinando o resultado final da classificação a respeito da amostra considerada. Em seguida, uma etapa de pós-processamento dinâmico, implementado através de uma rede recorrente é realizada. Esta permite mitigar os efeitos de oscilação que ocorre durante o reconhecimento de classes com características semelhantes. Nesta dissertação são investigados dois modelos de composição de especialistas. O primeiro é baseado no agrupamento de classes associadas a fonéticas semelhantes enquanto o segundo leva em conta a distribuição desequilibrada das amostras apresentadas nos dados de treinamento. A comparação do modelo proposto nesta dissertação com os trabalhos relacionados ao reconhecimento automático de fala indica um ganho de 7,62% em termos de acurácia.

Palavras-chave: Reconhecimento automático de fala. Reconhecimento fonético.
Redes neurais artificiais. Composição de especialistas.
Redes neurais recorrentes.