

RECONHECIMENTO DA FALA A PARTIR DE ESPECTROGRAMA USANDO MOBILENETV2

Projeto Demonstrativo 6 - Visão Computacional

Igor Bispo igor.rabbit99@gmail.com

Hevelyn Sthefany hevelyn.sthefany@gmail.com

December 5, 2018

Departamento de Ciência da Computação

Universidade de Brasília

Campus Darcy Ribeiro, Asa Norte

Brasília-DF, CEP 70910-900, Brazil,

Reconhecimento da Fala

Espectrograma

MobileNetV2

Resultados

Conclusão

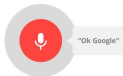
RECONHECIMENTO DA FALA

O reconhecimento de fala é uma tecnologia que procura identificar palavras faladas em um áudio e convertê-las em texto.





Hey Siri



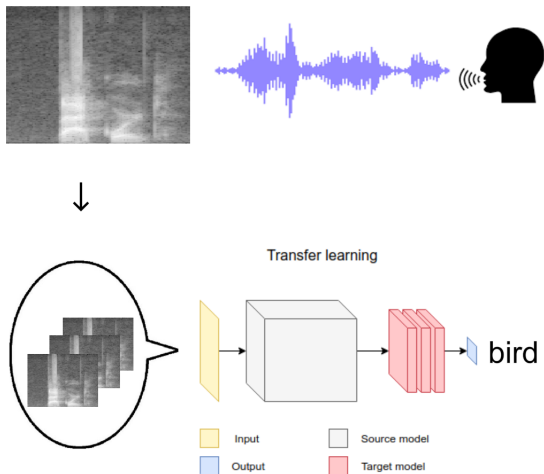
بنك أبو ظبي التجاري
ADCB
LONG LIVE AMBITION



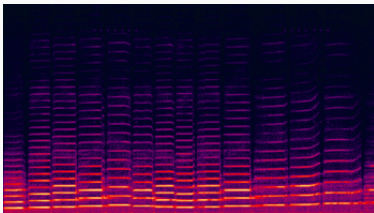
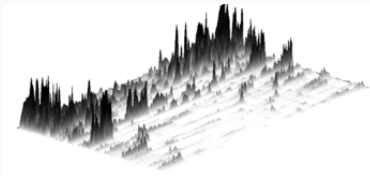
NUANCE

Siri, Open Mic+, Google Assistente, Siri são exemplos de aplicações que usam reconhecimento da fala.

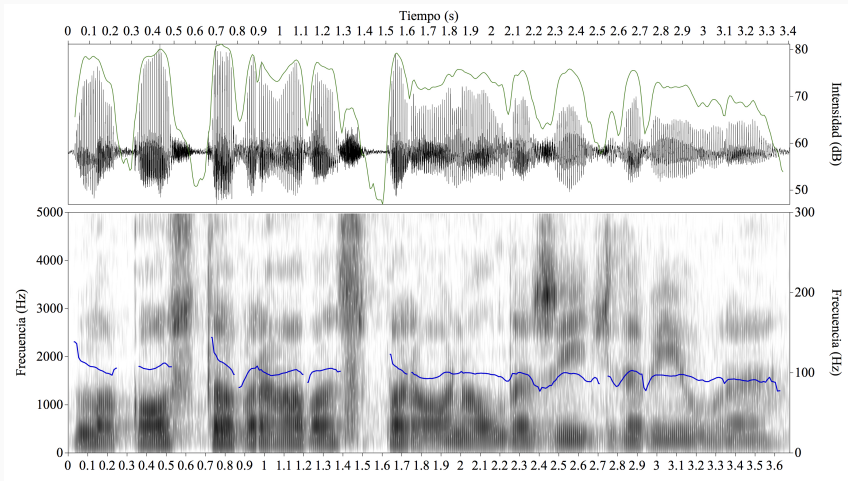
RECONHECIMENTO COM ESPECTROGRAMA USANDO MOBILENETV2



ESPECTROGRAMA



Gráficos que analisam dinamicamente a densidade espectral de energia. Os valores são indicados no plano **tempo X frequência** e cores indicam **intensidade** da densidade espectral de energia.



MOBILENETV2

Convoluções Separáveis em Profundidade Além de mapear uma única convolução em cada canal de entrada separadamente, adiciona a etapa da Convolução Pontual: Convolução com um tamanho de kernel de 1x1 que combina os recursos criados pela convolução em profundidade.

Resíduos Invertidos Blocos residuais conectam o início e o fim de um bloco convolucional com uma conexão de salto.

Gargalos Última convolução de um bloco residual tem uma saída linear antes de ser adicionada às ativações iniciais.

CONVULÇÕES SEPARÁVEIS

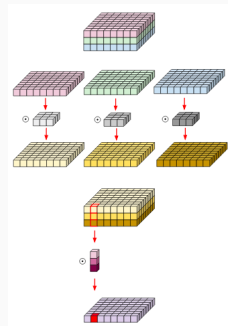
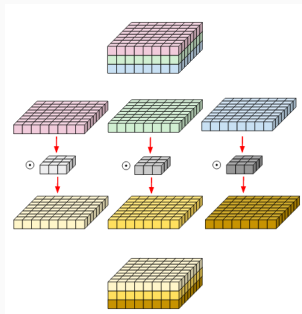
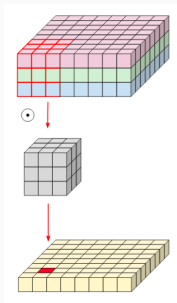


Figure: Na sequência: Convolução Normal, Convolução Profunda e Convolução Separável Profunda

RESÍDUOS INVERTIDOS

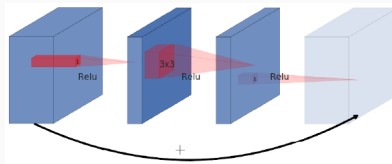


Figure: Um bloco residual conecta camadas largas com uma conexão de salto, enquanto as camadas intermediárias são estreitas

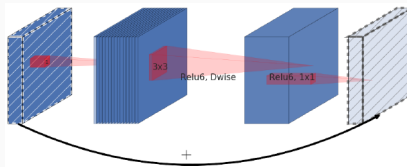
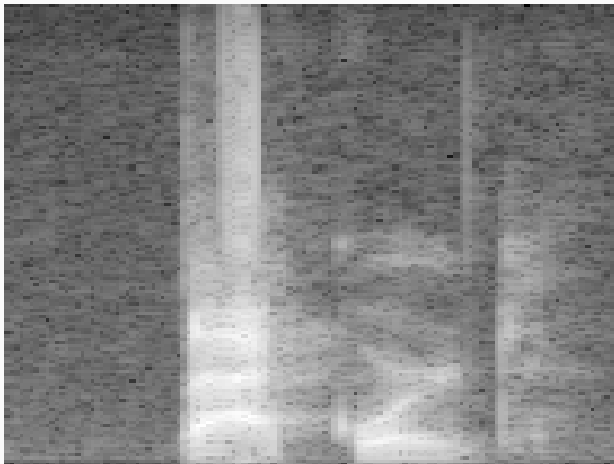


Figure: Um bloco residual invertido conecta camadas estreitas com uma conexão de salto enquanto as camadas intermediárias são largas

A última convolução de um bloco residual tem uma **saída linear** antes de ser adicionada às ativações iniciais. Simplesmente é descartado a última função de ativação do bloco convolucional.

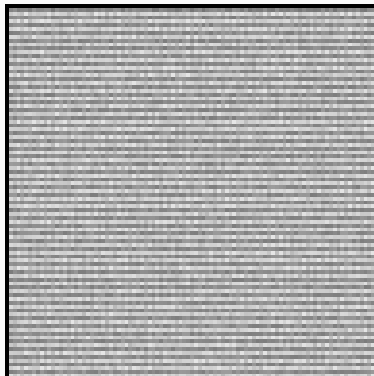
RESULTADOS

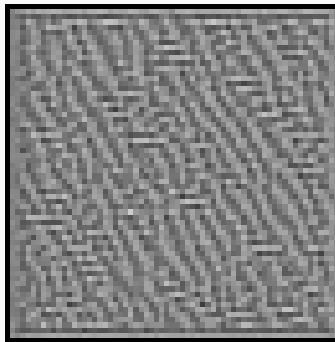
ESPECTROGRAMA "RIGHT"



Espectrograma classificado corretamente como "right" com 99% de confiança.

FILTRO APLICADO NA CONVOLUÇÃO PARA ESPECTROGRAMA





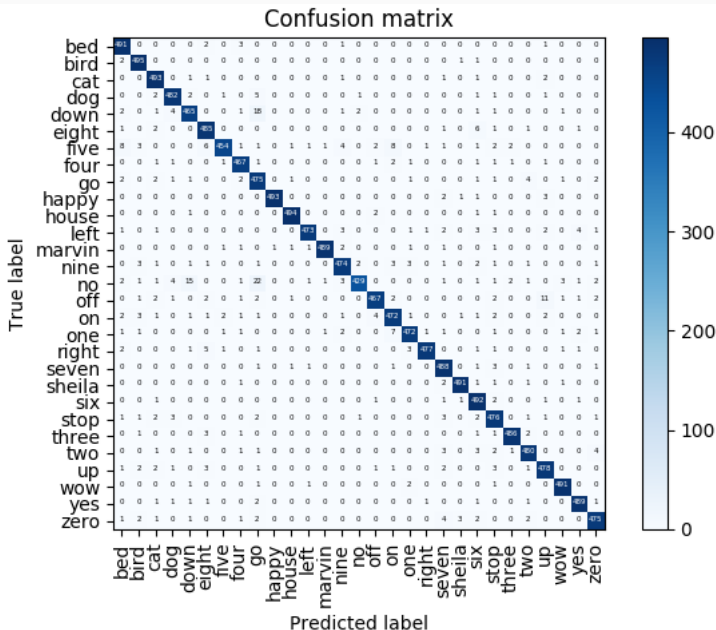
RELATÓRIO DE CLASSIFICAÇÃO

	precision	recall	f1-score	support
bed	0.95	0.98	0.97	500
bird	0.96	0.99	0.98	500
cat	0.96	0.99	0.97	500
dog	0.97	0.96	0.97	500
down	0.95	0.93	0.94	500
eight	0.95	0.97	0.96	500
five	0.98	0.91	0.94	500
four	0.97	0.93	0.95	500
go	0.88	0.95	0.92	500
happy	1.00	0.99	0.99	500
house	0.99	0.99	0.99	500
left	0.99	0.95	0.97	500
marvin	0.99	0.98	0.99	500
nine	0.96	0.95	0.96	500
no	0.99	0.86	0.92	500

RELATÓRIO DE CLASSIFICAÇÃO

off	0.98	0.93	0.96	500
on	0.95	0.94	0.95	500
one	0.97	0.94	0.96	500
right	0.99	0.95	0.97	500
seven	0.95	0.98	0.96	500
sheila	0.99	0.98	0.98	500
six	0.94	0.98	0.96	500
stop	0.94	0.95	0.94	500
three	0.99	0.97	0.98	500
two	0.97	0.96	0.96	500
up	0.95	0.96	0.95	500
wow	0.98	0.98	0.98	500
yes	0.98	0.98	0.98	500
zero	0.97	0.95	0.96	500
micro avg	0.97	0.92	0.96	14500
weighted avg	0.97	0.96	0.96	14500

MATRIZ DE CONFUSÃO - 10 EPOCHS



CONCLUSÃO

OBRIGADO!