



بهبود عملکرد دسته بندی صداهاى محیطی بر مبنای مدل یادگیری عمیق با ابعاد و پیچیدگی کم

چکیده

دسته بندی صداهاى محیطی یکی از موضوعات مهم در گستره مختلفی از کاربرد ها، (مثل شهر های هوشمند و نظارت صوتی) است، که به دلایل مختلفی از جمله الگوهای کمتر ساختار یافته، ایستایی کم، و تنوع زیاد درون- و بین-کلاسی، یکی از چالشی ترین حوزه های بازشناسی صوت است. بهبود های بدست آمده از روش های اخیر مبتنی بر یادگیری عمیق عمدتاً با افزایش عمق، پیچیدگی محاسباتی، و سایز شبکه همراه بوده است. در این رساله، ابتدا یک مدل مبتنی بر شبکه های عصبی کانولوشنال با سایز و پیچیدگی محاسباتی کم ارائه می کنیم که در طرح آن، ضمن الگوبرداری از سیستم شنوایی انسان، روش ادغام ویژگی عمومی نوینی را با توجه مشخصات زمانی و فرکانسی صداهاى محیطی و به نام ادغام تنک مناطق برجسته پیشنهاد می کنیم. روش ادغام ویژگی پیشنهادی با در نظر گرفتن تنوع بسیار زیاد الگوهای ورودی به عنوان اصلی ترین مانع برای یادگیری بهینه، مدل را با ایجاد یک گلوگاه اطلاعاتی به یادگیری مؤثر از الگوهای موجود در مناطق برجسته ورودی راهنمایی می کند. نتایج بدست آمده از ارزیابی ها نشان می دهند که مدل پیشنهادی با وجود پیچیدگی محاسباتی و سایز مدل بسیار کمتر از روش های سرآمد، دقت هایی قابل مقایسه با دقت های این روش ها هستند. همچنین، روش نوینی بر پایه انتقال یادگیری ارائه می کنیم که در آن روند مرسوم به روزرسانی را محدود به تنها زیرگروهی از نورون/کرنل ها می کنیم که بیشترین نقش را در عدم تشخیص صحیح نمونه ها دارند. ارزیابی های صورت گرفته بیانگر بهبود قابل ملاحظه در دقت تولید شده نسبت به روش مرسوم است.

دانشجو: **حامد ریاضتی سرشت**

استاد راهنما: **دکتر کریم محمدی**

اعضاء هیات داورى: **دکتر احمد آیت الهی؛ دکتر شهریار برادران شکوهی؛ دکتر محمد مهدی**

همایون پور؛ دکتر فرشاد الماس گنج