

موضوع:

Facial Emotion Recognition (FER)

عنوان مقاله:

A Survey of AI-based Facial Emotion Recognition: Features, ML & DL Techniques, Age-wise Datasets and Future Directions

خلاصه مقاله

این مقاله به بررسی جامع روش‌های تشخیص احساسات چهره (FER) مجموعه داده‌ها با تفکیک سنی، تکنیک‌های یادگیری ماشین و عمیق، چالش‌ها و روندهای آینده می‌پردازد. طبق تحقیقات، ۵۵ از انتقال احساسات از طریق چهره صورت می‌گیرد.

۱. مراحل پردازش در سیستم‌های FER

- **پیش‌پردازش (Pre-processing):** شامل تشخیص چهره و همترازی است. الگوریتم‌های محبوبی مانند Viola-Jones (سریع و کم‌هزینه) و MTCNN (دقیق‌تر با استفاده از شبکه‌های عصبی آشاری) استفاده می‌شوند. همچنین نرمال‌سازی نور و ژست صورت و افزایش داده‌ها (Data Augmentation) برای بهبود مدل انجام می‌شود.
- **استخراج ویژگی (Feature Extraction):** روش‌ها به دسته‌های بافت‌مبنا (مانند LBP و Gabor)، هندسی (مبتنی بر نقاط و فواصل اجزای صورت) و مبتنی بر یادگیری عمیق تقسیم می‌شوند.
- **طبقه‌بندی (Classification):** استفاده از الگوریتم‌هایی برای تشخیص نهایی احساس (مانند شادی، غم، خشم و غیره).

۲. مدل‌های یادگیری عمیق (Deep Learning)

- **CNN (شبکه‌های عصبی کانولوشن):** پرکاربردترین روش که شامل لایه‌های کانولوشن و پولینگ است. مدل‌های پیش‌آموزش‌دیده محبوبی مانند VGG، ResNet، (با اتصالات میانبر برای حل مشکل ناپدید شدن گرادیان) و Inception بررسی شده‌اند.
- **RNN/LSTM:** مناسب برای تحلیل داده‌های متوالی و ویدئویی جهت حفظ اطلاعات زمانی.
- **DBN (شبکه‌های باور عمیق):** مدلی مبتنی بر احتمالات که از ماشین‌های بولتزمن محدود شده (RBM) تشکیل شده است.
- **Autoencoder:** برای فشرده‌سازی و بازسازی ورودی و کاهش ابعاد استفاده می‌شود.

۳. بررسی مجموعه داده‌ها (Datasets) بر اساس سن

- **بزرگسالان:** بیشترین حجم داده را دارد. مجموعه داده‌های معروف شامل JAFFE، CK+ و FER2013 هستند.
- **کودکان:** داده‌های محدودتری دارد، مانند LIRIS (احساسات خودجوش)، CAFE و EmoReact.
- **سالمندان:** کمترین میزان پژوهش را دارد. مجموعه داده‌های اصلی شامل FACES و پایگاه داده تعاملات احساسی سالمندان است.

۴. چالش‌های پژوهشی

- انسداد (Occlusion): پوشانده شدن بخشی از صورت (مانند ماسک یا دست) که دقت را کاهش می‌دهد.
- عدم تعادل و سوگیری داده‌ها: اکثر داده‌ها از نظر نژاد، سن و جنسیت متعادل نیستند که منجر به عملکرد ضعیف مدل در دنیای واقعی می‌شود.
- داده‌های ۳بعدی و مادون قرمز: نیاز به استفاده بیشتر از داده‌های عمقی (D^3) و مادون قرمز برای غلبه بر مشکلات تغییر نور وجود دارد.

۵. روندهای آینده و کاربردها

- یادگیری انتقالی (Transfer Learning): استفاده از مدل‌های آموزش دیده روی داده‌های بزرگ برای مجموعه‌های کوچک.
- یادگیری فدرال (Federated Learning): آموزش غیرمتمرکز برای حفظ حریم خصوصی کاربران.
- هوش مصنوعی قابل توضیح (XAI): ارائه دلیل برای تصمیمات مدل جهت افزایش اعتماد.
- کاربردها: آموزش الکترونیکی (سنجش درک دانش آموز)، نظارت (تشخیص خستگی راننده)، پزشکی (کمک به کودکان اوتیستیک) و سرگرمی.