A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news

Khaled Obaid,Kuntara Pukthuanthong;JEF;2022

解读：王梦涵

2025 年 9 月 29 日

Motivation

- Investor sentiment helps a lot to understand and predict market returns
  - There has been lots of papers use news construct investor sentiment
- Photojournalism has increased in popularity due to modern technology and the demand for quick information
  - Photos in news can be attention-grabbing
  - Photos convey emotional info more effectively than words
- However, existing studies ignore *visual content* in news
- $\rightarrow$ How sentiment extracted from photos in the news media relates to market activities? How does it relates sentiment from text?

Question

- Q1: Whether *PhotoPes* in news relates to financial markets? If yes, how?
  - Yes, $PhotoPes_t \uparrow \Rightarrow MaketReturn_{t+1} \downarrow, MaketReturn_{t+5} \uparrow$
  - High/low *PhotoPes* can predict the next day abnormal trading volume' increase
- Q2: How pessimism embedded in photos and text interact?
  - *PhotoPes* and *TextPes* substitute for one another
- Q3: Difference between *PhotoPes* and *TextPes*
  - Q3-1: Whether photos can play an attention-grabbing role in newspapers?

    (Yes, Photos capture attention from text during periods when photos are salient)
  - Q3-2: Whether the relation between market returns and *PhotoPes* and *TextPes* during periods of elevated fear?

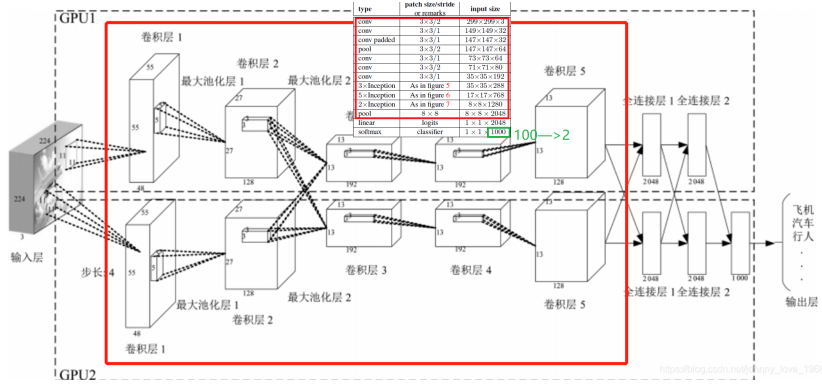
    (*PhotoPes*$_t$ doubles in high-fear periods; *TextPes* stays constant)

Contribution

- Literature on investor sentiment and news
  - Prior: Prove the **news text** can predict market returns and trading volume
  - Ext: Reveal that **news photos** contain content relevant to financial markets, and explore how does news photo content interact with news text content
- Literature on the psychology of visual stimuli in finance and economics.
  - Prior: Picture superiority effect *VS* Visual content is effective only with simple text
  - Ext: Reconcile the two kinds of literature
- Literature on the value of visual content in financial prediction
  - Prior: Political elections, loan decisions, firm market value...
  - Ext:Use news photos to proxy investor sentiment and predict returns
- Literature on large-scale photo classification

Data:*PhotoPes*–Google Inception (v3) model + transfer learning

- Transfer Learning: Use the DeepSent data set as training data for the linear layer(70%,10%,20%)



type	patch size/stride or remarks	input size
conv	3×3/2	299×299×3
conv	3×3/1	149×149×32
conv padded	3×3/1	147×147×32
pool	3×3/2	147×147×64
conv	3×3/1	73×73×64
conv	3×3/2	71×71×80
conv	3×3/1	35×35×192
3×Inception	As in figure 5	35×35×288
5×Inception	As in figure 6	17×17×768
2×Inception	As in figure 7	8×8×1280
pool	8×8	8×8×2048
linear	logits	1×1×2048
softmax	classifier	1×1×1000

- Randomly select 100 photos from our WSJ photo sample to test the model

Data:WSI and Market Return

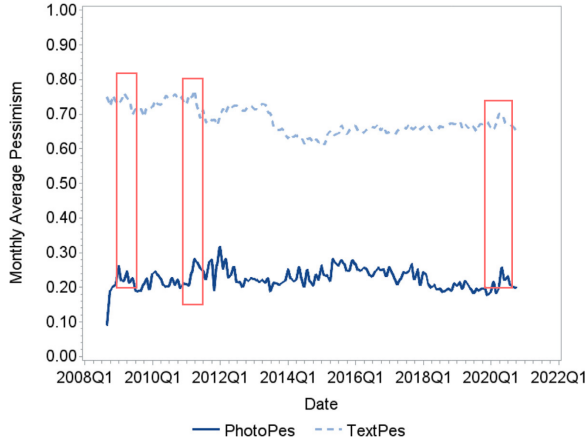
- WSJ:148,823 articles' headline, summary and photos (200809-202009)
  - $PhotoPes_t = \frac{\sum_i (Neg_{it})}{n_t}$
  - $TextPes_t = \frac{\sum_i (TextNeg_{it})}{n_t}$, where $TextNeg_{it}$ is the pessimism in each article evaluated by Stanford CoreNLP sentiment tool

Panel C: Correlations between sentiment variables

	PhotoPes
TextPes	0.079***
	<0.01

- Market Return: VWRETD;SPX;SPDR;INDU;DIA

Data: $PhotoPes_t$ and $TextPes_t$



- $PhotoPes_t$ and $TextPes_t$ move closer together during market turmoil

Regression Design for Q1 & Q2

- Q1: The impact of PhotoPes on market returns

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(R_t) + \beta_3 L5(R_t^2) + \beta_4 X_t + \varepsilon_t$$

- Q2: *PhotoPes* and sentiment embedded in text

PhotoPes' marginal effect: $\beta_1 + \beta_3 TextPes_{t-1}$;complement($\beta_3 < 0$),substitute$\beta_3 > 0$

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 (PhotoPes \times TextPes)_{t-1} +$$
$$\beta_4 L5(R_t) + \beta_5 L5(R_t^2) + \beta_6 X_t + \varepsilon_t,$$

  - R_t: daily log market returns
  - $L5$ transforms a variable into a row vector consisting of five lags of that variable
  - X_t: exogenous variables–an intercept, day-of-the-week indicators, $1_{RecessionPeriod}$

Q1: The impact of *PhotoPes* on market returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_{t-1}$	-0.042*	-1.837	-0.041*	-1.803	-0.040*	-1.787	-0.046**	-2.182	-0.047**	-2.183
$PhotoPes_{t-2}$	0.055**	2.004	0.051*	1.886	0.046*	1.726	0.043*	1.687	0.038	1.502
$PhotoPes_{t-3}$	-0.033	-1.324	-0.030	-1.213	-0.030	-1.294	-0.024	-1.053	-0.025	-1.142
$PhotoPes_{t-4}$	0.030	1.299	0.024	1.047	0.026	1.143	0.030	1.387	0.033	1.487
$PhotoPes_{t-5}$	0.057**	2.137	0.059**	2.228	0.056**	2.119	0.057**	2.193	0.054**	2.103
Sum t-1 to t-5	0.067		0.063		0.058		0.060		0.053	
Sum t-2 to t-5	0.109		0.104		0.098		0.106		0.100	
	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value	$\chi^2(1)$	p-value
$\chi^2(1)$[Sum t-1 to t-5 = 0]	2.272	0.132	2.081	0.149	1.700	0.192	1.979	0.160	1.644	0.200
$\chi^2(1)$[Sum t-2 to t-5 = 0]	6.615**	0.010	6.200**	0.013	5.466**	0.019	6.973***	0.008	6.257**	0.012
Adj. R-squared	0.033		0.038		0.029		0.042		0.040	
N	3044		3044		3044		3044		3044	

- *PhotoPes* has a non informational impact on market returns

Q1: The impact of *PhotoPes* on market returns–Channel

$$\bar{V}_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(|PhotoPes_t|) + \beta_3 L5(R_t) + \beta_4 L5(R_t^2) + \varepsilon_t.$$

$$\bar{V}_t = \beta L5(V_t) + \gamma X_t + \varepsilon_t$$

Variables	β	t-stat
$PhotoPes_{t-1}$	-0.003	-0.157
$PhotoPes_{t-2}$	-0.006	-0.278
$PhotoPes_{t-3}$	0.014	0.686
$PhotoPes_{t-4}$	-0.015	-0.786
$PhotoPes_{t-5}$	-0.011	-0.582
$\lvert PhotoPes_{t-1}\rvert$	0.080***	2.912
$\lvert PhotoPes_{t-2}\rvert$	0.013	0.453
$\lvert PhotoPes_{t-3}\rvert$	0.008	0.237
$\lvert PhotoPes_{t-4}\rvert$	0.059**	2.059
$\lvert PhotoPes_{t-5}\rvert$	0.015	0.511
Adj. *R*-squared	0.023	
N	3044	

- *PhotoPes* Changes investors belief and their behavior

Q2: PhotoPes and sentiment embedded in text

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_{t-1}$	-0.052**	-2.359	-0.049**	-2.229	-0.049**	-2.220	-0.054***	-2.600	-0.054***	-2.596
$TextPes_{t-1}$	-0.027	-0.816	-0.038	-1.183	-0.041	-1.241	-0.042	-1.390	-0.043	-1.372
$(PhotoPesxTextPes)_{t-1}$	0.038*	1.942	0.033*	1.754	0.034*	1.788	0.034*	1.899	0.034*	1.917
$PhotoPes_{t-2}$	0.056**	2.090	0.052**	1.980	0.048*	1.817	0.045*	1.801	0.040	1.595
$PhotoPes_{t-3}$	-0.027	-1.089	-0.025	-1.012	-0.025	-1.076	-0.020	-0.878	-0.021	-0.949
$PhotoPes_{t-4}$	0.032	1.394	0.026	1.125	0.028	1.228	0.031	1.454	0.034	1.581
$PhotoPes_{t-5}$	0.051*	1.936	0.053**	2.018	0.049*	1.896	0.051**	2.003	0.049*	1.897
$TextPes_{t-2}$	-0.040	-1.162	-0.045	-1.277	-0.042	-1.197	-0.047	-1.444	-0.042	-1.295
$TextPes_{t-3}$	-0.024	-0.640	-0.014	-0.387	-0.016	-0.432	-0.003	-0.078	-0.004	-0.119
$TextPes_{t-4}$	-0.016	-0.494	-0.019	-0.578	-0.020	-0.629	-0.014	-0.451	-0.021	-0.671
$TextPes_{t-5}$	0.090**	2.488	0.093***	2.603	0.095***	2.652	0.083**	2.497	0.089***	2.701

- *PhotoPes* and *TextPes* substitute for one another.

Q3: Regression Design for Q3

- Q3-1: Whether photos can play an attention-grabbing role in newspapers?

  E_t: $\mathbf{1}_{\text{day t is in the top or bottom decile of } \textit{PhotoPes}}$

- Q3-2: Whether the relation between market returns and *PhotoPes* and *TextPes* during periods of elevated fear?

  $F_t = \mathbf{1}_{\text{day t has an above-median fear score}}$

$$R_t = (F_t/E_t)\left[\beta_1 L5(\textit{PhotoPes}_t) + \beta_2 L5(\textit{TextPes}_t) + \beta_3(\textit{PhotoPes} \times \textit{TextPes})_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2)\right] + (1 - F_t/E_t)\left[\gamma_1 L5(\textit{PhotoPes}_t) + \gamma_2 L5(\textit{TextPes}_t) + \gamma_3(\textit{PhotoPes} \times \textit{TextPes})_{t-1} + \gamma_4 L5(R_t) + \gamma_5 L5(R_t^2)\right] + \beta_6 X_t + \varepsilon_t$$

Q3-1: Difference between *PhotoPes* and *TextPes*–Attention

	(1)				(2)				(3)			
	$VWRETD_t$				SPX_t				SPY_t			
	E_t=Salient photo period				E_t= Salient photo period				E_t= Salient photo period			
Variables	β	t-stat	γ	t-stat	β	t-stat	γ	t-stat	β	t-stat	γ	t-stat
$PhotoPes_{t-1}$	-0.070**	-2.479	-0.015	-0.332	-0.064**	-2.295	-0.016	-0.365	-0.063**	-2.260	-0.015	-0.342
$TextPes_{t-1}$	0.047	0.900	-0.070*	-1.883	0.031	0.606	-0.081**	-2.220	0.030	0.585	-0.080**	-2.137
$(PhotoPes \times TextPes)_{t-1}$	0.034	1.524	0.070	1.450	0.029	1.312	0.065	1.362	0.030	1.403	0.060	1.270
$PhotoPes_{t-2}$	0.100***	3.282	-0.034	-0.813	0.099***	3.316	-0.041	-0.978	0.094***	3.173	-0.044	-1.079
$PhotoPes_{t-3}$	-0.020	-0.659	-0.017	-0.411	-0.017	-0.565	-0.020	-0.498	-0.015	-0.526	-0.019	-0.489
$PhotoPes_{t-4}$	0.046*	1.710	0.009	0.216	0.042	1.571	-0.001	-0.033	0.042	1.632	0.003	0.066
$PhotoPes_{t-5}$	0.047	1.469	-0.009	-0.225	0.049	1.538	-0.004	-0.098	0.045	1.383	-0.006	-0.166
$TextPes_{t-2}$	0.044	0.684	-0.066*	-1.795	0.043	0.664	-0.071*	-1.912	0.051	0.761	-0.072*	-1.931
$TextPes_{t-3}$	-0.128**	-2.373	0.013	0.316	-0.113**	-2.174	0.022	0.538	-0.114**	-2.146	0.020	0.486
$TextPes_{t-4}$	-0.079	-1.412	0.002	0.057	-0.083	-1.488	-0.001	-0.018	-0.079	-1.488	-0.002	-0.050
$TextPes_{t-5}$	0.182***	3.133	0.059	1.520	0.186***	3.255	0.064	1.638	0.184***	3.135	0.065*	1.694

- Photos grab attention away from text during periods when photos are salient

Q3-2: Difference between *PhotoPes* and *TextPes*–High-fear Period

	(1)				(2)				(3)			
	$VWRETD_t$				SPX_t				SPY_t			
	F_t=Fear period				F_t=Fear period				F_t=Fear period			
Variables	β	t-stat	γ	t-stat	β	t-stat	γ	t-stat	β	t-stat	γ	t-stat
$PhotoPes_{t-1}$	-0.103**	-2.006	-0.037*	-1.893	-0.093*	-1.839	-0.036*	-1.822	-0.092*	-1.793	-0.036*	-1.848
$TextPes_{t-1}$	-0.018	-0.291	-0.019	-0.720	-0.031	-0.517	-0.025	-0.956	-0.035	-0.563	-0.026	-0.987
$(PhotoPes \times TextPes)_{t-1}$	0.087**	2.063	0.009	0.525	0.079*	1.920	0.005	0.294	0.076*	1.830	0.006	0.334
$PhotoPes_{t-2}$	0.065	1.100	0.050**	2.284	0.056	0.956	0.049**	2.272	0.053	0.879	0.049**	2.257
$PhotoPes_{t-3}$	0.018	0.357	-0.054**	-2.423	0.023	0.470	-0.054**	-2.455	0.018	0.374	-0.056**	-2.550
$PhotoPes_{t-4}$	0.093*	1.914	0.001	0.059	0.077	1.587	0.001	0.059	0.078*	1.646	0.004	0.194
$PhotoPes_{t-5}$	0.080	1.431	0.027	1.214	0.082	1.482	0.029	1.307	0.075	1.374	0.030	1.349
$TextPes_{t-2}$	-0.063	-0.919	-0.041	-1.510	-0.071	-1.031	-0.044*	-1.649	-0.059	-0.862	-0.046*	-1.697
$TextPes_{t-3}$	-0.079	-1.094	0.013	0.517	-0.062	-0.885	0.018	0.721	-0.065	-0.920	0.019	0.756
$TextPes_{t-4}$	-0.062	-0.937	0.007	0.263	-0.069	-1.046	0.006	0.235	-0.066	-1.022	0.007	0.267
$TextPes_{t-5}$	0.134*	1.916	0.061**	2.122	0.138**	2.006	0.062**	2.161	0.138**	2.025	0.061**	2.150

- Photos being more effective than text at conveying fear or traumatic events

Trading strategies

Panel A: Summary statistics of trading strategies

Strategy	*N*	*Mean*	*t-stat*	*Std dev*	*SR*
PhotoPes	1992	0.058	3.251	1.119	0.052
TextPes	1891	0.037	2.085	1.166	0.032
Combined	1221	0.054	3.547	0.980	0.055
Index	3034	0.047	2.246	1.325	0.036

Panel B: Time series regression

	(1)		(2)	
	Combined strategy$_t$		*PhotoPesstrategy*$_t$	
Variables	β	*t*-stat	β	*t*-stat
Alpha	0.021*	1.742	0.014	1.302
Mkt_Rf_t	51.0***	13.347	69.4***	22.912
SMB_t	-1.830	-0.317	-6.612	-1.418
HML_t	-15.6***	-3.517	-13.9***	-3.799
UMD_t	-4.947*	-1.743	-5.489**	-2.121
ST_Rev_t	6.754*	1.708	3.252	1.002
Adj. *R*-squared	0.545		0.706	
N	3034		3034	

New ideas

- Inappropriate:
  - Control vars:macro economic vars; News' ‘‘dissemination features (e.g., views, shares, dwell time)...
  - *TextPes*：Distinguish different news topic
- Other possible data source
  - Video; Voice; Memes; GIF
  - Social Media